Locally Self-Similar Processes and Their Wavelet Analysis

Joseph E. Cavanaugh
Department of Statistics, University of Missouri, Columbia, MO 65211, USA
Yazhen Wang
Department of Statistics, University of Connecticut, Storrs, CT 06269, USA
J. Wade Davis
Department of Statistics, University of Missouri, Columbia, MO 65211, USA

1. Introduction

A stochastic process $Y(t)$ is defined as *self-similar* with self-similarity parameter H if for any positive stretching factor c, the distribution of the rescaled and reindexed process $c^{-H}\,Y(c\,t)$ is equivalent to that of the original process $Y(t)$. This means for any sequence of time points $t_1, \ldots, t_n$ and any positive constant c, the collections $\{c^{-H}\,Y(c\,t_1), \ldots, c^{-H}\,Y(ct_n)\}$ and $\{Y(t_1), \ldots, Y(t_n)\}$ are governed by the same probability law. As a consequence, the qualitative features of a sample path of a self-similar process are invariant to magnification or shrinkage, so that the path will retain the same general appearance regardless of the distance from which it is observed.

Although self-similar processes were first introduced in a theoretical context by Kolmogorov (1941), statisticians were made aware of the practical applicability of such processes through the work of B.B. Mandelbrot (Mandelbrot and van Ness, 1968; Mandelbrot and Wallis, 1968, 1969). Self-similarity is a pervasive characteristic in naturally occurring phenomena. As a result, self-similar processes have been used to successfully model data arising in a variety of different scientific fields, including hydrology, geophysics, medicine, genetics, economics, and computer science. Recent applications include Buldyrev et al. (1993), Lelend et al. (1994), Ossadnik et al. (1994), Percival and Guttorp (1994), Peng et al. (1992, 1995a, 1995b), and Willinger et al. (1998).

The dynamics of a self-similar process $Y(t)$ are principally dictated by the value of the self-similarity parameter or scaling exponent H. For practical applications, it may be assumed that $Y(t)$ has finite second moments, and that its associated increment process $X(t) = Y(t) - Y(t-1)$ is stationary. Under these assumptions, H may be taken to lie over the interval $(0, 1)$, and the value of H may be used in describing the autocorrelation structure of the increment sequence $X(t), X(t+1), \ldots$ (see Beran, 1994, pp. 52–53). For $H \in (1/2, 1)$, $X(t)$ is characterized by serial correlations that decay slowly, and therefore exhibits long-range dependence or long memory. For $H \in (0, 1/2)$, $X(t)$ is characterized by serial correlations that decay rapidly and sum to zero. For $H = 1/2$, $X(t)$ is serially uncorrelated. The estimation of H as a constant has been extensively studied, predominantly in the context of long memory where it is assumed that $H \in (1/2, 1)$. A partial list of relevant references includes Geweke and Porter-Hudak (1983), Taylor and Taylor (1991), Constantine and Hall (1994), Chen et al. (1995), Robinson (1995), Taqqu et al. (1995), Abry and Sellan (1996), Comte (1996), McCoy and Walden (1996), Hall et al. (1997), Kent and Wood (1997), Moulines and Soulier (1999), Abry et al. (2000), Bardet et al. (2000), Abry et al. (2001), and Bardet et al. (2001).

In modeling applications, treating the self-similarity parameter H as a constant implies that the self-similar features of the underlying phenomenon persist over time. However, many phenomena exhibit self-similar patterns that change as the phenomenon itself evolves: this characteristic may be embodied in the associated data by sections of irregular roughness. To adequately model such data, the class of self-similar processes must be expanded to allow the scaling exponent to vary as a function of time. Moreover, a procedure must be available that provides a statistical characterization of the exponent's progression.

In what follows, we consider a class of processes that are locally (as opposed to globally) self-similar. The defining property of this class is based on the covariance function, and relates the local autocorrelation of the process to the value of $H(t)$. We provide two examples of stochastic processes that are locally self-similar. One of these, generalized fractional Brownian motion, is explored in detail. We then propose and describe a procedure based on wavelets for constructing an estimate of the time-varying scaling exponent of a locally self-similar process. We establish a consistency result for this estimate. We investigate the

effectiveness of the procedure in a simulation study, and demonstrate its applicability in several practical analyses.

Techniques for estimating a constant self-similarity parameter H are often based on log-linear regression (e.g., Geweke and Porter-Hudak, 1983; Taylor and Taylor, 1991; Constantine and Hall, 1994). Such methods frequently exploit an approximate log-linear relationship between either the spectrum of $X(t)$ or the variogram of $Y(t)$ and the time index t, using least-squares regression to obtain the estimate of H. With a locally self-similar process $Y(t)$, $H(t)$ is a function of t; as a result, the associated increment process $X(t)$ is non-stationary. To estimate $H(t)$ locally, we take advantage of an approximate local log-linear relationship between the square of the wavelet transformation of $Y(t)$ and the scale for the transformation. Local least-squares regression is used to obtain the estimate of $H(t)$. The basic procedure is detailed in Wang et al. (2001) and was first presented by Wang et al. (1997). The method may be viewed as an extension of previously proposed log-linear regression techniques for the estimation of constant H.

Replacing the role of the spectrum or variogram with that of the wavelet transformation is logical for the problem at hand. Since wavelets are ideal for characterizing changes in scale content over time, they can accommodate the time-varying dynamics of $Y(t)$ induced by changes in $H(t)$. Our reliance on wavelets is further motivated by the computational efficiency of the discrete wavelet transformation as well as the successful application of wavelets in problems dealing with both non-stationarity and self-similarity (e.g., Farge et al., 1993; Hwang and Mallat, 1994). In fact, for globally self-similar processes, log-linear regression methods that utilize the wavelet transformation have been recently proposed for the estimation of constant H (e.g., Abry et al., 2000; Bardet et al., 2000; Abry et al., 2001). An early inspiration for these methods as well as for our procedure is provided by an illustration in Daubechies (1992, pp. 301–303) for estimating the Hölderian exponent.

The exposition in this chapter is organized as follows. Section 2 defines and briefly discusses locally self-similar processes. Section 3 provides an in–depth exploration of the properties of generalized fractional Brownian motion. Section 4 presents the procedure for the estimation for $H(t)$ via the continuous wavelet transformation. Section 5 outlines the implementation of the procedure to data collected in discrete time via the discrete wavelet

transformation. Section 6 presents simulations designed to check the effectiveness of the procedure, and section 7 features three practical applications. Section 8 concludes.

2. Locally self-similar processes

Let $Y(t)$ represent a mean-zero stochastic process with covariance function

$$\Sigma_t(u_1, u_2) = \mathrm{Cov}\{Y(t+u_1), Y(t+u_2)\} \tag{2.1}$$

and variogram

$$\Upsilon_t(u_1, u_2) = \mathrm{Var}\{Y(t+u_1) - Y(t+u_2)\}. \tag{2.2}$$

The process $Y(t)$ is said to be *locally self-similar* if

$$\Sigma_t(u_1, u_2) = \Sigma_t(0,0) - C(t)\,|u_1 - u_2|^{2\,H(t)}\{1 + o(1)\},\ \text{as}\ |u_1| + |u_2| \to 0, \tag{2.3}$$

where $C(t)$ is a non-negative function of t, and $H(t)$ represents the local scaling exponent function, or scaling function for short. Note that (2.3) suggests

$$\Upsilon_t(u_1, u_2) = 2\,C(t)\,|u_1 - u_2|^{2\,H(t)}\{1 + o(1)\},\ \text{as}\ |u_1| + |u_2| \to 0. \tag{2.4}$$

Provided that

$$\Sigma_t(u_1, u_1) + \Sigma_t(u_2, u_2) = 2\Sigma_t(0,0) + o(|u_1 - u_2|^{2\,H(t)}),\ \text{as}\ |u_1| + |u_2| \to 0,$$

the relations (2.3) and (2.4) are equivalent.

Definition (2.3) relates the local autocorrelation of $Y(t)$ to the value of $H(t)$. With $C(t) > 0$, for fixed $|u_1 - u_2|$ near 0, the covariance between $Y(t+u_1)$ and $Y(t+u_2)$ becomes smaller as $H(t)$ moves from 1 towards 0. As a consequence, sections of sample paths where $H(t)$ is near 0 should appear more rough and erratic than sections where $H(t)$ is near 1.

Let $X(t) = Y(t) - Y(t-1)$ denote the increment process of $Y(t)$, with covariance function

$$\Psi_t(u_1, u_2) = \mathrm{Cov}\{X(t+u_1), X(t+u_2)\}.$$

The *Wigner-Ville distribution* of $X(t)$, which functions as a local version of the power spectrum, is defined as

$$g_t(\lambda) = \frac{1}{2\pi} \int \Psi_t(\tau/2, -\tau/2)\, e^{-i\,\lambda\,\tau} d\tau.$$

It can be shown that

$$g_t(\lambda) = K(t)\,\lambda^{1-2\,H(t)}\,\{1 + o(1)\}, \quad \text{as } \lambda \to 0, \tag{2.5}$$

where $K(t)$ is a non-negative function of t. When $H(t) \in (1/2, 1)$, (2.5) implies that $g_t(\lambda)$ is unbounded at $\lambda = 0$. This spectral characteristic is indicative of long-range dependence.

If the increment process $X(t)$ is stationary, $g_t(\lambda)$ is independent of t and reduces to the ordinary power spectrum of $X(t)$. In this instance, $K(t)$ and $H(t)$ are constant, and (2.5) becomes

$$g(\lambda) = K\lambda^{1-2\,H}\,\{1 + o(1)\}, \quad \text{as } \lambda \to 0. \tag{2.6}$$

Estimation of constant H in (2.6) has been considered by many authors, mostly in the context of long-memory time series where $H \in (1/2, 1)$.

We now present and discuss examples of processes that exhibit locally self-similar behavior. First, consider the process defined for real $t \geq 0$ by the stochastic integral

$$Y(t) = \int_{-\infty}^{0} \left\{(t-u)^{H(t)-1/2} - (-u)^{H(t)-1/2}\right\} dW(u) + \int_0^t (t-u)^{H(t)-1/2}\, dW(u), \tag{2.7}$$

where $W(t)$ is standard Brownian motion, and $H(t) \in (0, 1)$ represents the scaling function. This process is an extension of fractional Brownian motion, fBm, that allows for the self-similarity parameter to vary over time. For this reason, the process is called *generalized fractional Brownian motion,* gfBm.

In the following section, we explore many of the important theoretical and conceptual properties of gfBm. We provide simple, explicit forms for its variance function, covariance function, and variogram. We show that gfBm may be locally approximated by fBm, and establish that gfBM obeys the law of the iterated algorithm. Finally, we derive the fractal, Hausdorff, and local dimensions of gfBM sample paths.

The increment process of fBm is called *fractional Gaussian noise*, fGn. fGn is stationary with a spectrum that satisfies (2.6). The increment process of gfBm is non-stationary (unless $H(t)$ is constant) with a spectrum that satisfies (2.5). Since gfBm relaxes many homogeneous restrictions of fBm, the process may be used to model many natural phenomena with time-varying dynamics that cannot be accommodated by either standard Brownian motion or fBm (cf. Mandelbrot, 1983; Gonçalvès and Flandrin, 1993; Flandrin and Gonçalvès, 1994).

Next, consider the process defined for non-negative integer t as

$$\Phi(\mathrm{B})(1-\mathrm{B})^{H(t)-1/2}X(t) = \Theta(\mathrm{B})\epsilon(t), \tag{2.8}$$

where B is a backshift operator defined by $\mathrm{B}X(t) = X(t-1)$, $\Phi(\mathrm{B})$ and $\Theta(\mathrm{B})$ are polynomials in B having characteristic roots outside the unit circle, $\epsilon(t)$ is Gaussian white noise, and $H(t) \in (0,1)$ represents the scaling function. This process is an extension of a fractional autoregressive integrated moving-average, or fARIMA, process. Again, the extension allows for the self-similarity parameter to evolve over time. We refer to this process as a *generalized fARIMA* or gfARIMA process.

A fARIMA process is stationary with a spectrum that satisfies (2.6). A gfARIMA process is non-stationary (unless $H(t)$ is constant) with a spectrum that satisfies (2.5). It can be shown that the normalized partial sums of a fARFIMA process have the same limiting distribution as that of a globally self-similar process (see Beran, 1994, pp. 48–50). Analogously, it can be shown that the normalized partial sums of a gfARIMA process have the same limiting distribution as that of a locally self-similar process. Thus, a fARIMA process may be regarded as the increment process for a globally self-similar process, and a gfARIMA process may be regarded as the increment process for a locally self-similar process.

fARIMA processes have been extensively studied in the context of long-memory time series (see Beran, 1994). gfARIMA processes have even greater potential for widespread applicability, since they provide for the modeling of non-stationary time series that exhibit irregular patterns of roughness.

A special case of (2.8) is defined by

$$(1-\mathrm{B})^{H(t)-1/2}X(t) = \epsilon(t). \tag{2.9}$$

We refer to such a process as *generalized fractionally integrated noise*, gfin, since it represents an extension of fractionally integrated noise, fin.

Many of the fundamental properties of the two aforementioned locally self-similar processes have yet to be established. In the next section, we investigate the first of these processes, gfBm. We exhibit the form of its variance function, covariance function, and variogram. We show that gfBm may be locally approximated by fBm, and establish that

gfBM obeys the law of the iterated algorithm. We also derive the fractal, Hausdorff, and local dimensions of gfBM sample paths.

3. Generalized fractional Brownian motion

Fractional Brownian motion has been used in modeling a host of natural phenomena from a wide range of scientific fields (e.g., Mandelbrot and van Ness, 1968; Mandelbrot, 1983; Peitgen et al., 1992; Wornell and Oppenheim, 1992). Such a process may be defined by (2.7) provided that $H(t) \equiv H$ is regarded as constant. As with any self-similar process, the process dynamics are largely characterized by the value of $H \in (0, 1)$. When $H = 1/2$, (2.7) reduces to ordinary Brownian motion. When $H > 1/2$, (2.7) involves fractional integration, which is a smoothing operation. In this case, the process is said to be *persistent,* yielding sample paths that appear trend-like, exhibiting gradual, sustained ripples. When $H < 1/2$, (2.7) involves fractional differentiation, which is an unsmoothing operation. In this case, the process is said to be *anti-persistent,* yielding sample paths that appear noisy, exhibiting jagged, erratic shifts.

Like Brownian motion, fractional Brownian motion has many properties that are homogeneous in nature, including stationary increments, global self-similarity, and constant local dimensions. Such properties prohibit fBm from being used to describe complicated phenomena characterized by different modes of regularity. For instance, in modeling landscapes via random fractals and $1/f$ signals, processes that allow for varying patterns of roughness are desirable (cf. Mandelbrot, 1983; Wang, 1997). In this section, we explore the properties of generalized fractional Brownian motion and demonstrate that the process has the capability to describe the dynamics of complex phenomena.

We begin in subsection 3.1 by exhibiting the form of the covariance function for gfBm, which then leads to expressions for the variance function and the variogram. In subsection 3.2, we investigate local properties of gfBm. We show that the process may be locally approximated by fBm, and establish that it obeys the law of the iterated algorithm. Finally, in subsection 3.3, we derive the fractal, Hausdorff, and local dimensions of its sample paths.

3.1 Covariance function, variance function, variogram

In the definition of gfBm provided in section 2, we assume that the process $Y(t)$ has mean zero. Thus, the covariance between $Y(t)$ and $Y(s)$ is given by $E\{Y(t)\,Y(s)\}$ and the variance of $Y(t)-Y(s)$ by $E[\{Y(t)-Y(s)\}^2]$. In this section, we will denote the covariance function and variogram respectively by writing

$$R(s,t) = E\{Y(t)\,Y(s)\} \qquad \text{and} \qquad V(s,t) = E[\{Y(t)-Y(s)\}^2].$$

Thus, in the notation of (2.1) and (2.2), we have $\Sigma_t(u_1,u_2) = R(t+u_1, t+u_2)$ and $\Upsilon_t(u_1,u_2) = V(t+u_1, t+u_2)$

We note that in the definition of gfBm in section 2, a constant starting value $Y(0)$ may be added to the right-hand side of (2.7). Without loss of generality, we have assumed this constant is zero.

The following theorem provides an explicit form for the covariance function of gfBm. All proofs appear in the Appendix.

THEOREM 1.

$$\begin{aligned} R(s,t) \;=\; & t^{H(s)+H(t)}\left[\{H(s)+H(t)\}^{-1} + \int_0^\infty f_H(u;s,t)\,du\right]/2 \\ & + s^{H(s)+H(t)}\left[\{H(s)+H(t)\}^{-1} + \int_0^\infty f_H(u;t,s)\,du\right]/2 \\ & - |t-s|^{H(s)+H(t)}\left[\{H(s)+H(t)\}^{-1} + \int_0^\infty f_H(u;s\wedge t, s\vee t)\,du\right]/2, \end{aligned}$$

where

$$f_H(u;s,t) \;=\; u^{H(s)+H(t)-1} + (1+u)^{H(s)+H(t)-1} - 2\,u^{H(s)-1/2}\,(1+u)^{H(t)-1/2}.$$

From Theorem 1, formulas immediately follow for the variance function and the variogram of gfBm.

COROLLARY 1.

$$\mathrm{Var}\{Y(t)\} = t^{2\,H(t)}\left[\{2\,H(t)\}^{-1} + \int_0^\infty f_H(u;t,t)\,du\right],$$

$$
\begin{aligned}
\mathrm{V}(t,s) &= |t-s|^{H(s)+H(t)}\left[\{H(s)+H(t)\}^{-1}+\int_0^\infty f_H(u;s\wedge t, s\vee t)\,du\right] \\
&\quad + t^{2H(t)}\left[\{2H(t)\}^{-1}+\int_0^\infty f_H(u;t,t)\,du\right] \\
&\quad - t^{H(s)+H(t)}\left[\{H(s)+H(t)\}^{-1}+\int_0^\infty f_H(u;s,t)\,du\right] \\
&\quad + s^{2H(s)}\left[\{2H(s)\}^{-1}+\int_0^\infty f_H(u;s,s)\,du\right] \\
&\quad - s^{H(s)+H(t)}\left[\{H(s)+H(t)\}^{-1}+\int_0^\infty f_H(u;t,s)\,du\right].
\end{aligned}
$$

REMARK 1.1. If $H(t)\equiv H$ is constant, $Y(t)$ is fBm with self-similarity parameter H. The formula in Theorem 1 reduces to the exact covariance function of fBm. Indeed, since

$$
\begin{aligned}
f_H(u;s,t) &= u^{2H-1}+(1+u)^{2H-1}-2\,u^{H-1/2}\,(1+u)^{H-1/2} \\
&= \left\{u^{H-1/2}-(1+u)^{H-1/2}\right\}^2,
\end{aligned}
$$

we have

$$
R(s,t) = \left(t^{2H}+s^{2H}-|t-s|^{2H}\right)\left[(2H)^{-1}+\int_0^\infty\left\{u^{H-1/2}-(1+u)^{H-1/2}\right\}^2 du\right]/2.
$$

REMARK 1.2. If $H(s)+H(t)=1$, then

$$
H(s)-1/2 = 1/2-H(t) = \{H(s)-H(t)\}/2,
$$

and

$$
f_H(u;s,t) = 2-2\left(\frac{u}{1+u}\right)^{\frac{H(s)-H(t)}{2}}.
$$

The formula for $R(s,t)$ simplifies to

$$
R(s,t) = (s\wedge t)\left[1+\int_0^\infty\left\{2-\left(\frac{u}{1+u}\right)^{\frac{H(s)-H(t)}{2}}-\left(\frac{u}{1+u}\right)^{\frac{H(t)-H(s)}{2}}\right\}du\right].
$$

As an example of an instance where the preceding formula would be applicable, suppose $H(t)=\alpha$ for $t\in(0,1/3]$ and $H(t)=1-\alpha$ for $t\in[2/3,1)$. Then over the intervals $(0,1/3]$ and $[2/3,1)$, $Y(t)$ corresponds to fBm with respective self-similarity parameters α and $1-\alpha$, and for $s\in(0,1/3]$ and $t\in[2/3,1)$,

$$
R(s,t) = s\left[1+\int_0^\infty\left\{2-\left(\frac{u}{1+u}\right)^{\alpha-1/2}-\left(\frac{u}{1+u}\right)^{1/2-\alpha}\right\}du\right].
$$

For $\alpha \neq 1/2$, α and $1-\alpha$ fall on opposite sides of $1/2$, so $Y(t)$ is persistent over one of the intervals ($H > 1/2$) and anti-persistent over the other ($H < 1/2$). Moreover, the preceding covariance function provides the covariance between the persistent and anti-persistent sections of the process.

3.2 Local properties

Consider a fixed time point t_*. The following theorem shows for small δ, the variance of $Y(t_* + \delta) - Y(t_*)$ is approximately the same as that for fBm with self-similarity parameter $H(t_*)$. The theorem further establishes that $Y(t)$ may be locally approximated at t_* by fBm.

THEOREM 2. *Suppose $H(t)$ is twice continuously differentiable at t_*.*

(1)

$$V(t_* + \delta, t_*) = |\delta|^{2\,H(t_*)} \left[\{2\,H(t_*)\}^{-1} + \int_0^\infty f_H(u; t_*, t_*)\, du \right] + O\left(\delta^2\right).$$

(2) *Let $\tilde{Y}(t)$ be fBm with parameter $H(t_*)$, i.e.,*

$$\tilde{Y}(t) = \int_{-\infty}^0 \left\{ (t-u)^{H(t_*)-1/2} - (-u)^{H(t_*)-1/2} \right\} dW(u) + \int_0^t (t-u)^{H(t_*)-1/2}\, dW(u).$$

Then,

$$\mathrm{Var}\{Y(t_* + \delta) - \tilde{Y}(t_* + \delta)\} = O\left(\delta^2\right).$$

REMARK 2.1. If $H(t_*) < 1/2$, under the assumption that $H(t)$ is once continuously differentiable, the proof of Theorem 2 can be modified to show that the results of the theorem hold with error terms of order δ.

REMARK 2.2. Note that the first result of Theorem 2 reflects the local self-similarity of gfBm, since it clearly suggests the variogram property (2.4).

The following theorem establishes the law of the iterated logarithm for gfBm.

THEOREM 3. *Let*

$$M(t_*, \delta) = \sup \{ |Y(s) - Y(t)|; s, t \in (t_* - \delta, t_* + \delta) \}.$$

Then there exist constants c and C depending only on $H(t_*)$ such that

$$\liminf_{\delta \to 0} \frac{M(t_*, \delta)}{(\delta / \log\,|\log\,\delta|)^{H(t_*)}} = c, \qquad \limsup_{\delta \to 0} \frac{M(t_*, \delta)}{\delta^{H(t_*)} / (\log\,|\log\,\delta|)^{1/2}} = C.$$

3.3 Fractal, Hausdorff, and local dimensions

Fundamental characteristics of the behavior of a stochastic process are quantified via the fractal, Hausdorff, and local dimensions of its sample paths. In what follows, we show that for gfBm, these dimensions are solely determined by $H(t)$. The results reinforce the assertion that the dynamics of a locally self-similar process $Y(t)$ are principally dictated by the scaling function $H(t)$.

We begin by providing general definitions for the fractal, Hausdorff, and local dimensions based on the graph of a function $g(x)$. The first two of these definitions will involve the graph of the function over an interval $[a, b] \subset [0, \infty)$, denoted by $G = \{(x, g(x)); x \in [a, b]\} \subset \mathbb{R}^2$.

Let $d(z, \epsilon)$ represent a disk centered at $z \in \mathbb{R}^2$ with radius $\epsilon > 0$, and let $N_\epsilon(G)$ be the minimum number of disks $d(z, \epsilon)$ that cover G. The *fractal dimension* (or *box dimension*) of $g(x)$ on $[a, b]$ is defined as

$$\Delta_{[a,b]}(g) = \lim_{\epsilon \to 0} \frac{\log N_\epsilon(G)}{|\log \epsilon|}. \tag{3.1}$$

Next, let $\{d_i\}$ denote a collection of open disks with diameter bounded by $\epsilon > 0$ chosen so that the union of the disks covers G. Let the *α-dimensional outer measure* be defined as

$$S_{\alpha,\epsilon}(G) = \inf \left\{ \sum_i (\text{diameter } d_i)^\alpha \right\},$$

where the infimum is taken over all such collections $\{d_i\}$, and the sum is taken over all disks in a collection. The *Hausdorff dimension* of $g(x)$ on $[a, b]$ is defined as

$$D_{[a,b]}(g) = \inf\{\alpha : S_\alpha(G) = 0\} = \sup\{\alpha : S_\alpha(G) = \infty\}. \tag{3.2}$$

(See Mandelbrot, 1983; Adler, 1981; Barnsley, 1993; Tricot, 1995.)

Finally, the *local dimensions* of $g(x)$ at a point x are defined in terms of (3.1) and (3.2) as

$$\Delta_x(g) = \lim_{\delta \to 0} \Delta_{[x-\delta, x+\delta]}(g), \qquad D_x(g) = \lim_{\delta \to 0} D_{[x-\delta, x+\delta]}(g).$$

The following two theorems provide the fractal and Hausdorff dimensions for the sample paths of gfBm.

THEOREM 4. *Let $H_{ab} = \min\{H(t);\, t \in [a, b]\}$. If $H_{ab} < 1/2$, suppose $H(t)$ is once continuously differentiable on $[a, b]$; if $H_{ab} \geq 1/2$, suppose $H(t)$ is twice continuously differentiable*

on $[a, b]$. *Then the fractal and Hausdorff dimensions of the sample paths of* $Y(t)$ *are given by*

$$\Delta_{[a,b]}(Y) = D_{[a,b]}(Y) = 2 - H_{ab}.$$

REMARK 4.1. Theorem 4 confirms the intuition that the fractal and Hausdorff dimensions of gfBm are dominated by the roughest segments of its sample paths.

Our final theorem provides the local dimensions for the sample paths of gfBm. The result is obtained by applying Theorem 4 to the interval $[a, b] = [t - \delta, t + \delta]$ and letting $\delta \to 0$.

THEOREM 5. *If* $H(t) < 1/2$, *suppose* $H(t)$ *is once continuously differentiable at* t; *if* $H_{ab} \geq 1/2$, *suppose* $H(t)$ *is twice continuously differentiable at* t. *Then the local dimensions of the sample paths of* $Y(t)$ *are given by*

$$\Delta_t(Y) = D_t(Y) = 2 - H(t).$$

4. Estimating the local self-similarity parameter

The results of subsection 3.3 demonstrate the importance of the role of the scaling function $H(t)$ in characterizing the behavior of a locally self-similar process. In this section, we outline the procedure used to estimate this function.

Our development is based on the application of the continuous wavelet transformation to a locally self-similar process $Y(t)$ defined in continuous time. Of course, in practice, $Y(t)$ will be measured over a collection of discrete time points, yielding a finite sample of measurements. Thus, in section 5, we discuss the implementation of the procedure to a finite series via the discrete wavelet transformation.

To begin, let ψ denote the mother wavelet, assumed to have at least two vanishing moments and to be at least twice continuously differentiable. Let $TY(a, t)$ denote the wavelet transformation of the locally self-similar process $Y(t)$ corresponding to the scale a and the location t. We may then write

$$TY(a,t) = a^{-1/2} \int \psi\left(\frac{u-t}{a}\right) Y(u)\, du = a^{1/2} \int \psi(x)\, Y(t + a\, x)\, dx. \tag{4.1}$$

Using (4.1), (2.3), and the first vanishing moment of ψ, we obtain

$$\begin{aligned} E\left\{|TY(a,t)|^2\right\} &= \frac{1}{a}\int\int \psi\left(\frac{u-t}{a}\right)\,\psi\left(\frac{v-t}{a}\right)\,E\{Y(u)\,Y(v)\}\,du\,dv \\ &= a\int\int \psi(x)\,\psi(y)\,E\{Y(t+a\,x)\,Y(t+a\,y)\}\,dx\,dy \\ &= a\int\int \psi(x)\,\psi(y)\,\Sigma_t(a\,x,a\,y)\,dx\,dy \\ &\sim a\int\int \psi(x)\,\psi(y)\,\left\{\Sigma_t(0,0)-C(t)\,|a\,x-a\,y|^{2\,H(t)}\right\}\,dx\,dy \\ &= C_1\,a^{1+2\,H(t)}, \qquad \text{as } a\to 0, \end{aligned} \tag{4.2}$$

where

$$C_1 = -C(t)\int\int |x-y|^{2\,H(t)}\psi(x)\,\psi(y)\,dx\,dy.$$

Now let

$$y_t(a) = \log\left\{|TY(a,t)|^2\right\},$$

$$C_2 = E\left(\log\left[|TY(a,t)|^2\Big/\,E\left\{|TY(a,t)|^2\right\}\right]\right),$$

and

$$\varepsilon_t(a) = \log\left[|TY(a,t)|^2\Big/\,E\left\{|TY(a,t)|^2\right\}\right] - C_2.$$

Then clearly,

$$y_t(a) = C_2 + \log\left[E\left\{|TY(a,t)|^2\right\}\right] + \varepsilon_t(a). \tag{4.3}$$

Note that (4.2) and (4.3) imply the approximate regression model

$$y_t(a) \approx c + \{2\,H(t)+1\}\,\log\,a + \varepsilon_t(a), \qquad \text{for small scale } a, \tag{4.4}$$

where $c = \log\,C_1 + C_2$.

Assuming that the error terms in model (4.4) are at least approximately uncorrelated, the model suggests the use of ordinary least squares to estimate $H(t)$. The general procedure is outlined as follows.

1. Select a sequence of small scales $a_1 > \cdots > a_k$, say $a_j = 2^{-j}$ where $j = 1,\ldots,k$.

2. Define a set of bivariate data (x_j, y_j), $j = 1,\cdots,k$, by setting

$$x_j = \log\,a_j \qquad \text{and} \qquad y_j = y_t(a_j) \qquad \text{for each } j.$$

3. Evaluate the least-squares estimate of $H(t)$ in (4.4) via

$$\hat{H}(t) = \left\{ \frac{\sum (x_j - \bar{x})\ (y_j - \bar{y})}{\sum (x_j - \bar{x})^2} - 1 \right\} /2, \tag{4.5}$$

where $\bar{x} = \sum x_j/k$, $\bar{y} = \sum y_j/k$.

In the Appendix, we establish the following consistency result for the estimator $\hat{H}(t)$.

THEOREM 6. *Suppose that Y(t) is Gaussian with a covariance function that satisfies (2.3). Then as $k \to \infty$, $\hat{H}(t)$ converges in probability to $H(t)$.*

Intuitively, the magnitude of $TY(a,t)$ reflects the content of scale a present in the process at time t. The approximate regression model (4.4) implies that at time t, the degree to which the magnitude of $TY(a,t)$ decreases with a is governed by $H(t)$. According to the model, the larger the value of $H(t)$, the greater the rate at which the scale content decays with the corresponding scale. Since fine scale content should be more prevalent at times where a sample path appears rough than at times where a path appears smooth, larger values of $H(t)$ should therefore be associated with smoother sample paths.

Having conceptualized the procedure in terms of the continuous wavelet transformation, we now discuss its implementation to a finite series via the discrete wavelet transformation.

5. Implementation of the estimation procedure

In practice, we generally observe a process $Y(t)$ at a set of discrete time points. For convenience, we will assume that the points are scaled to lie over the interval [0,1). We will also assume that the points are equally spaced, and that the sample size is a power of two. Thus, we may index our measurements with $t_i = (i-1)/n$, where $i = 1, \ldots, n = 2^J$.

If the sample size is not a power of two, the series may be extended to achieve such a sample size by repeating the final data value, and the results may be adjusted accordingly. (See the application in subsection 7.1 for an example.)

Let the vector $Y = [Y(t_1), \ldots, Y(t_n)]'$ represent our sample of n measurements on $Y(t)$. To implement the estimation procedure, the discrete wavelet transformation or DWT will be applied to the vector Y, and an estimate of $H(t)$ will be derived from the transformed data.

The DWT of Y is a one-to-one transformation which reconfigures the information in Y so that two objectives are served: (1) large-scale behavior (e.g., trends, long-term cycles) may be easily delineated from small-scale behavior (e.g., local irregularities, noise), (2) the evolution of scale content may be monitored over time. The discrete Fourier transformation or DFT serves the first objective yet not the second, since the DFT is formulated under the assumption that frequencies persist over time.

The DWT may be written in the form $\mathcal{W}Y$, where $\mathcal{W}$ is an $n \times n$ orthogonal matrix that depends on both the wavelet and the boundary adjustment (Cohen et al., 1993a; Daubechies, 1994). Fast algorithms with complexity of order n are available for performing both the DWT and the inversion of the transformation that results in the reconstruction of the original data. (Such algorithms are even less computationally intensive than the fast Fourier transformation, which has complexity of order $n \log_2 n$.) Thus, the DWT provides an efficient, information-preserving mechanism for the characterization of scale content over time.

To label the $n = 2^J$ coefficients of the DWT, we index $n-1$ of the coefficients dyadically:

$$y_{j,k}; \quad k = 0, \ldots, 2^j - 1; \quad j = 0, \ldots, J-1.$$

We then label the remaining coefficient $y_{-1,0}$. The coefficient $y_{j,k}$ is referred to as the DWT of Y at level j and location $k2^{-j}$. For $k = 0, \ldots, 2^j - 1$ and $j = 0, \ldots, J-1$, the quantity $y_{j,k}$ may be viewed as an approximation to $TY(2^{-j}, k2^{-j})$, the continuous wavelet transformation $TY(a,t)$ evaluated at scale $a = 2^{-j}$ and location $t = k2^{-j}$ (Donoho and Johnstone, 1994). Intuitively, the magnitude of $y_{j,k}$ reflects the content of scale $a = 2^{-j}$ in that portion of the series which lies over the interval $[k2^{-j}, (k+1)2^{-j})$.

In what follows, we outline the implementation of the estimation procedure for $H(t)$ in a series of five steps.

1. Perform the DWT on the series Y to obtain the DWT coefficients

$$\{y_{j,k} \mid k = 0, \ldots, 2^j - 1; j = 0, \ldots, J-1\}.$$

2. Partition the sampling interval $[0,1)$ into 2^l nonoverlapping subintervals of equal length, where l is an integer chosen such that $0 \le l \le (J-1)$. The 2^l subintervals are of the

form

$$I_m = [m2^{-l}, (m+1)2^{-l}); \quad m = 0, \ldots, 2^l - 1.$$

3. Choose an integer $J' \leq (J-1)$ such that $2^{-J'}$ represents the largest scale to be used in the procedure. All DWT coefficients $y_{j,k}$ where $J' \leq j \leq (J-1)$ will be used in estimating the $H(t)$ curve.

4. For each I_m, $m = 0, \ldots, 2^l - 1$, pool together values of $y_{j,k}$ to be used in constructing the local estimate for $H(t)$. Use the $y_{j,k}$ corresponding to levels j such that $J' \leq j \leq (J-1)$, and to locations $k2^{-j}$ such that the associated interval $[k2^{-j}, (k+1)2^{-j})$ overlaps with I_m. Specifically, for each $m = 0, \ldots, 2^l - 1$, perform the following steps.

   a. Define the bivariate collection of data

$$\begin{aligned}(X_m, Y_m) \quad = \quad & \left\{ \left(\log\left(2^{-j}\right), \log\left(|y_{j,k}|^2\right) \right) \quad | \right. \\ & \left. [k2^{-j}, (k+1)2^{-j}) \cap I_m \neq \phi; \;\; 0 \leq k \leq 2^j - 1, \;\; J' \leq j \leq (J-1) \right\}.\end{aligned}$$

   b. Fit an ordinary least-squares line to (X_m, Y_m), treating the X_m as the regressor measurements and the Y_m as the response measurements.

   c. Solve for the estimate $\hat{H}(t)$ by adjusting the estimate of the slope in the least-squares fit by first subtracting 1 and then dividing by 2, as indicated in (4.5).

   One may envision each $\hat{H}(t)$ as estimating the average value of the scaling function $H(t)$ over the corresponding subinterval I_m. The appropriate time index for the $\hat{H}(t)$ associated with I_m might be regarded as the midpoint of I_m, namely $2^{-l-1}(2m+1)$.

5. Construct a curve from the collection of local estimates $\hat{H}(t)$ by employing a smoothing algorithm, such as local polynomial smoothing. This curve then serves to approximate the shape of the scaling function $H(t)$.

As mentioned in the introduction, if a globally self-similar process $Y(t)$ has stationary increments and finite second moments, H may be taken to lie over the interval $(0, 1)$. Values of H outside of this range lead to processes that are not of practical interest. However, in

applying our estimation procedure for $H(t)$ to real data, estimates of $H(t)$ that fall below 0 or above 1 will occasionally occur. Such estimates may arise for a variety of reasons. For instance, if the procedure is applied to the first difference of a locally self-similar process, estimates of $H(t)$ between -1 and 0 generally result; if the procedure is applied to the partial sums of a locally self-similar process, estimates of $H(t)$ between $+1$ and $+2$ generally result. Thus, estimates of $H(t)$ outside the interval (0,1) may imply that the underlying series (or subseries) must be suitably differenced or accumulated in order to obtain a series amenable to our algorithm (i.e., one that may be regarded as locally self-similar with locally stationary increments and finite second moments). Of course, such estimates may also occur if the underlying series (or subseries) reflects the dynamics of a process that lies outside the realm of locally self-similar processes. This issue, however, is beyond the scope of the present paper.

6. Simulations

We test the performance of our algorithm in twelve simulation sets: six based on the gfBm process (2.7) and six based on the partial sums of the gfin process (2.9). We simulate the processes so that the time index t is confined to the interval $[0, 1)$.

To simulate realizations of the gfBm process (2.7), consider setting

$$W_n(t_i) = \frac{1}{\sqrt{n}} \sum_{k=1}^{i} \epsilon(t_k); \quad t_i = \frac{(i-1)}{n};\ i = 1, \ldots, n; \tag{6.1}$$

where the $\epsilon(t_k)$ are variates of a Gaussian white noise process. For large n, the values of $W_n(t_i)$ may be treated as realizations of Brownian motion, since the normalized partial sum in (6.1) approximates the stochastic integral that defines such a process. The second of the two stochastic integrals that comprise (2.7) may then be approximated by its discretized analogue

$$\begin{aligned} Y_{2,n}(t_i) &= \sum_{k=1}^{i} (t_i - t_k)^{H(t_i)-1/2} \{W_n(t_k) - W_n(t_{k-1})\} \qquad \{W(t_0) \equiv 0\} \\ &= \frac{1}{\sqrt{n}} \sum_{k=1}^{i} (t_i - t_k)^{H(t_i)-1/2} \epsilon(t_k); \quad t_i = \frac{(i-1)}{n};\ i = 1, \ldots, n. \end{aligned}$$

Similarly, the first of the two stochastic integrals in (2.7) can be approximated by a sum of the form

$$Y_{1,N}(t_i) = \frac{1}{\sqrt{N}} \sum_{k=-N}^{0} \left\{ (t_i - u_k)^{H(t_i)-1/2} - (-u_k)^{H(t_i)-1/2} \right\} \epsilon(u_k);$$
$$u_k = \left(\frac{k}{N} \right) K; \; k = 0, -1, \ldots, -N; \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n;$$

where the $\epsilon(u_k)$ are variates of a Gaussian white noise process, and the integer $N > 0$ and constant $K > 0$ are chosen such that $(u_{-N}, u_{-N+1}], (u_{-N+1}, u_{-N+2}], \ldots, (u_{-1}, u_0]$ provides a sufficiently fine partition of a suitably large interval $(-K, 0]$. The sequence

$$Y_n(t_i) = Y_{1,N}(t_i) + Y_{2,n}(t_i); \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n;$$

may then be treated as realizations of the process (2.7).

We remark that the contribution of $Y_{1,N}(t_i)$ to $Y_n(t_i)$ is negligible for the purpose at hand. This contribution is therefore omitted in generating our sample paths. We also note that certain components in the sums that comprise $Y_{1,N}(t_i)$ and $Y_{2,n}(t_i)$ may be of the form 0^a where $a \leq 0$. Components of the form 0^0 can be treated as 1, and those of the form 0^a where $a < 0$ can be treated as 0.

To simulate realizations of the gfin process defined by (2.9), we use the fact that (2.9) may be represented as an infinite moving average of the form

$$X(t_i) = \sum_{k=0}^{\infty} a(t_i, k) \epsilon(t_{i-k}), \tag{6.2}$$

where the $\epsilon(t_{i-k})$ are variates of a Gaussian white noise process, and the $a(t_i, k)$ are defined by

$$a(t_i, k) = \frac{\Gamma(k + \{H(t_i) - 1/2\})}{\Gamma(k+1)\Gamma(H(t_i) - 1/2)} \tag{6.3}$$

(see Beran, 1994, p. 65). (Here, $\Gamma(\cdot)$ denotes the gamma function.) The infinite sum in (6.2) may be approximated by the finite sum

$$X_n(t_i) = \sum_{k=0}^{N} a(t_i, k) \epsilon(t_{i-k}); \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n; \tag{6.4}$$

provided that N is chosen to be a suitably large integer. The sequence (6.4) may then be treated as realizations of the process (6.2), or equivalently, (2.9). Accordingly, for large n,

the normalized partial sums

$$Y_n(t_i) = \frac{1}{\sqrt{n}} \sum_{k=1}^{i} X_n(t_k); \quad t_i = \frac{(i-1)}{n};\ i = 1, \ldots, n;$$

may be regarded as realizations of a locally self-similar process with a covariance function of the form (2.3). These sums are therefore amenable to our algorithm.

In computing the coefficients $a(t_i, k)$, we note that evaluating $\Gamma(k)$ for large k may lead to overflow. This problem can be avoided by using the log gamma function to compute $\log a(t_i, k)$ and exponentiating. We also note that when $H(t_i) = 1/2$, formula (6.3) for the $a(t_i, k)$ involves $\Gamma(0)$. In this instance, the $a(t_i, k)$ should be assigned the value 1 when $k = 0$ and 0 otherwise.

We generate six samples of the gfBm process (2.7) and six of the gfin process (2.9) using the following six scaling functions:

$$\begin{aligned}
H_1(t) &= 0.7 && \text{(constant)} \\
H_2(t) &= 0.2 + 0.5t && \text{(linear)} \\
H_3(t) &= 3.0(t - 0.5)^2 + 0.1 && \text{(quadratic)} \\
H_4(t) &= 4.0\{(t - 0.5)^3 + 0.125\} && \text{(cubic)} \\
H_5(t) &= 0.25\{\log(10.0\, t + 0.5) + 1.0\} && \text{(logarithmic)} \\
H_6(t) &= 0.01 \exp(4.0\, t) + 0.25 && \text{(exponential)}
\end{aligned}$$

In each instance, we consider a sample of size $n = 2^{12} = 4096$ $(J = 12)$.

In the estimation algorithm, we use the least asymmetric wavelet with a width of 8 (Daubechies, 1992) and the interval boundary conditions (Cohen et al., 1993b). We partition the sampling interval $[0, 1)$ into $32 = 2^5$ $(l = 5)$ subintervals, which results in 32 local estimates of $H(t)$ corresponding to the time indices $t = 1/64, 3/64, \ldots, 63/64$. In pooling values of $y_{j,k}$ for each estimate, we use the coefficients corresponding to the scales $2^{-4}, 2^{-5}, \ldots, 2^{-11}$ (i.e., $J' = 4$). We smooth the estimates using local polynomial smoothing.

For the simulation sets based on the gfBm processes, the sample paths and the estimated curves for $H(t)$ are displayed in Figs. 1.1 through 1.12; for the sets based on the gfin processes, the partial sums for the sample paths and the estimated curves for $H(t)$ are displayed in Figs. 2.1 through 2.12.

Each estimated curve effectively approximates the general shape of the corresponding scaling function $H(t)$. Note that the curves reflect a minor amount of negative bias in the $\hat{H}(t)$: this phenomenon is also apparent in the simulation results reported by Wang et al. (2001). Not surprisingly, the number of subintervals used in the sampling interval partition (i.e., the choice of l) affects the bias: the larger the number of subintervals, the smaller the bias. Of course, employing a finer partition also inflates the variability of the estimated curve. However, the impact of this increased variation can be partly controlled by the manner in which the estimate is smoothed.

The encouraging results obtained in our simulations suggest that our estimation procedure should result in an effective characterization of $H(t)$ in large-sample settings. We now apply the procedure to three real data sets: two from hydrology and one from computer science.

7. Applications

7.1 Vertical ocean shear measurements

Percival and Guttorp (1994) analyze a set of vertical ocean shear measurements. The data for the measurements are collected by dropping a probe into the ocean that records the water velocity every 0.1 meter as it descends. The "time" index is depth (in meters). The shear measurements (in 1/seconds) are obtained by taking a first difference of the velocity readings over 10 meter intervals, and then applying a low-pass filter to the differenced readings.

Vertical shear measurements display characteristics typical of self-similar processes; in particular, the increments of such a series often exhibit long-memory behavior. The data considered by Percival and Guttorp consist of 6875 values collected from a depth of 350.0 meters down to 1037.4 meters. The authors analyze 4096 of the values (chosen from the middle of the series) using wavelets and the Allan variance. Their justification for selecting the central part of the sample for their analysis is that "this subseries can be regarded as a portion of one realization of a process whose first backward difference is a stationary process" (p. 334). In other words, this part of the sample can be regarded as a sample path of a globally self-similar process with stationary increments.

We analyze the entire sample under the premise that the series can be treated as a realization of a locally self-similar process. Our goal is to estimate the scaling function $H(t)$.

Rather than redefining the time index so that the sampling interval is $[0, 1)$, we retain the values of the original index (in meters). Since the size of the series is not a power of two, we extend its length to $2^{13} = 8192$ ($J = 13$) by repeating the last value in the series 1317 times. We then apply our estimation algorithm by partitioning the sampling interval into $32 = 2^5$ ($l = 5$) subintervals, which results in 32 local estimates of $H(t)$. Only the first 27 estimates are relevant, since the remaining estimates pertain to the augmented portion of the series. We smooth these 27 estimates using local polynomial smoothing.

In implementing our procedure, we employ the least asymmetric wavelet with a width of 8 and the interval boundary conditions. We use the DWT coefficients corresponding to the scales $2^{-4}, 2^{-5}, \ldots, 2^{-12}$ (i.e., $J' = 4$).

The series is plotted in Fig. 3.1, and the smoothed estimated curve for $H(t)$ in Fig. 3.2. In both figures, the range of t corresponding to the portion of the sample analyzed by Percival and Guttorp is delineated with dotted lines. We note that the $H(t)$ curve provides strong evidence that the self-similarity parameter is not constant. The curve varies over the range from 0.65 to 1.00, which reinforces the notion that the increments of the series exhibit long-memory behavior. However, the shape of the curve indicates that the nature of the long-range dependence changes as the depth increases. To model $H(t)$ as a constant over the entire sampling interval would be to ignore the local aspect of self-similarity that appears to characterize the series.

The estimated curve for $H(t)$ appears somewhat quadratic. The shape of the curve is similar to that featured in Wang et al. (2001), where the series is analyzed using the same estimation algorithm, albeit with a different choice for the wavelet. Note that the curve appears reasonably flat over the middle section between the dotted lines. This supports the argument of Percival and Guttorp that the central section of the series may be regarded as globally self-similar with stationary increments. However, when taken as a whole, the series exhibits depth-varying self-similarity patterns that can only be characterized by modeling the scaling exponent as a function of t and by estimating $H(t)$ accordingly.

7.2 Yearly minimum water levels for the Nile River

In the study of long-range dependence, one of the most widely analyzed time series is a collection of yearly minimum water levels for the Nile River. The 660 levels were recorded from 622 A.D. to 1281 A.D. at the Roda Gauge near Cairo, Egypt. Percival and Walden (2000, pp. 190–193) and Beran (1994, p. 22) discuss the data and provide interesting perspectives on its history.

Beran (1994, pp. 117–118) initially analyzes the series using two models: fGn (the increment process of fBm), and fin. Using a discrete version of Whittle's estimator, he constructs point estimates as well as 95% confidence intervals for the global self-similarity parameter H. The results are summarized below.

Model	Point Estimate	95% Confidence Interval
fGn	0.84	(0.79,0.89)
fin	0.90	(0.84,0.96)

Beran (1994, p. 206-207) presents a second analysis of the series based on a suspected change point in the self-similarity parameter around 722 A.D. He partitions the initial 600 observations of the series into 6 sections consisting of 100 observations each. The Whittle estimates for each of these 6 sections are featured below.

Time Interval	(622,721)	(722,821)	(822,921)	(922,1021)	(1022,1121)	(1122,1221)
Point Estimate	0.54	0.85	0.86	0.83	0.84	0.93

Treating the series as a realization of the increment process of a locally self-similar process, we analyze the initial $2^9 = 512$ $(J = 9)$ measurements. The objective of our analysis is to estimate $H(t)$ over two sections of the series: the points collected prior to 722 A.D., and the points collected from 722 onwards. We will then compare our results to those obtained by Beran.

To make the series amenable to our algorithm, the mean of the observations is subtracted, and the normalized partial sums of the resulting data are computed. The original series is plotted in Fig. 4.1 and the partial sums in Fig. 4.2. The time point of interest, 722 A.D., is demarcated with a dotted line.

Rather than redefining the time index so that the sampling interval is $[0, 1)$, we retain the values of the original index (in years). We apply our algorithm by partitioning the sampling interval into $16 = 2^4$ ($l = 4$) subintervals, which results in 16 local estimates of $H(t)$. Of these 16 estimates, the first 3 apply to the time period prior to 722 A.D., and the remaining 13 apply to the period from 722 onwards. Using local polynomial smoothing, we smooth the first 3 estimates and the last 16 estimates separately, allowing for a discontinuity in the $H(t)$ curve at the time point 722 A.D.

In implementing our procedure, we employ the least asymmetric wavelet with a width of 12 and the interval boundary conditions. We use the DWT coefficients corresponding to the scales $2^{-5}, 2^{-6}, \ldots, 2^{-8}$ (i.e., $J' = 5$).

The smoothed estimated curve for $H(t)$ is featured in Fig. 4.3, again with the time point 722 A.D. highlighted with a dotted line. The first 5 of Beran's 6 Whittle estimates are superimposed on the plot with dots. Clearly, our results support Beran's assertion that a change point in $H(t)$ occurs around 722 A.D.

Over the section of the series extending from 722 A.D. onwards, our results seem quite consistent with Beran's Whittle estimates: the $H(t)$ curve appears roughly level and oscillates around 0.85. In fact, the mean of the 13 local estimates of $H(t)$ is 0.85; curiously, this value lies within both of Beran's confidence intervals for the global estimation of H.

For the section of the series collected prior to 722 A.D., our results are somewhat different from Beran's. Here, the $H(t)$ curve rises abruptly from about 0.1 to about 0.4, and the mean of the 3 local estimates of $H(t)$ is 0.25. These results are indicative of anti-persistent behavior. On the other hand, Beran's Whittle estimate for this section is 0.54, which implies that the observations are characteristic of white noise.

As a point of interest, Percival and Walden (2000, pp. 326–327, 386–388) discuss a suspected shift in the variance of the series around 715 A.D. They speculate that this shift is due to a change in the method used to measure the minimum water level. A test is conducted to determine if a change point occurs at 715 A.D., and the test shows significance. Based on the final 512 points in the series, a maximum likelihood estimate is constructed for a long-memory parameter δ, which is related to the self-similarity parameter H by $H = \delta + 0.5$. The result is $\hat{\delta} = 0.4452$ (p. 286), which corresponds to the estimate $\hat{H} = 0.9452$. Thus, for

the latter section of the series, the analysis of Percival and Walden indicates long-range dependence which is slightly stronger than that reflected by either our analysis or the analyses of Beran.

7.3 Ethernet network traffic measurements

Our final analysis considers a series comprised of $2^{18} = 262,144$ Ethernet local area network (LAN) traffic measurements collected over a one-hour period at the Bellcore Morris Research and Engineering Center. Each observation represents the number of packets sent over the Ethernet during a 10 millisecond period. A larger version of the series, featuring 360,000 observations spanning the entire hour of the experiment, is analyzed by Leland et al. (1994). The data were collected during a period of "normal" traffic load.

As explained by Leland et al., self-similarity is present in Ethernet LAN traffic data due to the absence of a natural length of a "burst." At various time scales, ranging from a few milliseconds to minutes to hours, traffic bursts are present that are similar in structure. Thus, a series of traffic measurements has the same basic appearance whether it is viewed over very short time scales (e.g., milliseconds), or over much longer scales (e.g., hundreds of seconds).

Analyzing the series of length 360,000, Leland et al. recovers the self-similarity parameter H in three ways: by using the slope of a variance-time curve, by using the slope of an R/S plot, and by using a maximum likelihood estimator based on the periodogram. The results are as follows.

Method	Point Estimate	95% Confidence Interval
slope of variance-time curve	0.87	NA
slope of R/S plot	0.90	NA
periodogram-based MLE	0.90	(0.85,0.95)

We consider an analysis of the series of length $2^{18} = 262,144$ that treats the series as a realization of the increment process of a locally self-similar process. Our objective is to estimate $H(t)$, and to interpret our results in light of the analysis of Leland et al.

To make the series amenable to our algorithm, the mean of the observations is subtracted, and the normalized partial sums of the resulting data are computed. The first 1000 observations of the original series (covering a 10 second interval) are plotted in Fig. 5.1; the entire set of partial sums (covering roughly 2600 seconds) is plotted in Fig. 5.2.

Rather than redefining the time index so that the sampling interval is $[0, 1)$, we retain the values of the original index (in 10 millisecond increments). We apply our algorithm by partitioning the sampling interval into $1024 = 2^{10}$ ($l = 10$) subintervals, which results in 1024 local estimates of $H(t)$. We smooth these estimates using local polynomial smoothing.

In implementing our procedure, we employ the least asymmetric wavelet with a width of 8 and the interval boundary conditions. We use the DWT coefficients corresponding to the scales $2^{-4}, 2^{-5}, \ldots, 2^{-9}$ (i.e., $J' = 4$).

The smoothed estimated curve for $H(t)$ is featured in Fig. 5.3. The limits of the aforementioned confidence interval reported by Leland et al. are represented in the figure by dotted lines. Note that the curve for $H(t)$ is somewhat flat, and lies entirely within the confidence interval limits. Both our analysis and that of Leland et al. indicate that the series exhibits strong long-range dependence.

Our analysis suggests that in applications where the self-similarity parameter is roughly constant, our procedure should produce an estimated $H(t)$ curve which accurately reflects this characteristic. The simulation results based on the constant scaling function $H_1(t)$ reinforce this assertion. Thus, in analyzing data where it is not apparent whether the self-similar behavior remains globally constant or changes over time, our procedure provides a promising alternative to methods that assume the former.

8. Conclusion

Many naturally occurring phenomena produce data that exhibits self-similar behavior which evolves as the phenomenon progresses. Adequate modeling of such data requires the consideration of locally self-similar processes with time-varying scaling exponents.

In the preceding chapter, we have considered two examples of locally self-similar processes: generalized fractional Brownian motion, gfBm, and the limit of the normalized partial sums of generalized fractional ARIMA processes, gfARIMA. We have investigated several key

theoretical properties of gfBM. Additionally, we have proposed, discussed, and investigated an algorithm for estimating the time-varying scaling function $H(t)$ of a locally self-similar process. Our algorithm exploits the time-scale localization facility of wavelets to produce a consistent estimator of $H(t)$ at a given time point t. Our simulation results, which feature both gfBm and gfARIMA processes, indicate that our method provides an accurate reflection of the progression of $H(t)$. Moreover, our applications illustrate that our method can be used to quantify time-dependent self-similarity patterns that arise in actual spatial and temporal series.

Figure Captions

Fig. 1.1. Sample path of gfBm with scaling function $H_1(t)$ ($n = 4096$).

Fig. 1.2. Estimate of $H_1(t)$ for gfBm sample path in Fig. 1.1.

Fig. 1.3. Sample path of gfBm with scaling function $H_2(t)$ ($n = 4096$).

Fig. 1.4. Estimate of $H_2(t)$ for gfBm sample path in Fig. 1.3.

Fig. 1.5. Sample path of gfBm with scaling function $H_3(t)$ ($n = 4096$).

Fig. 1.6. Estimate of $H_3(t)$ for gfBm sample path in Fig. 1.5.

Fig. 1.7. Sample path of gfBm with scaling function $H_4(t)$ ($n = 4096$).

Fig. 1.8. Estimate of $H_4(t)$ for gfBm sample path in Fig. 1.7.

Fig. 1.9. Sample path of gfBm with scaling function $H_5(t)$ ($n = 4096$).

Fig. 1.10. Estimate of $H_5(t)$ for gfBm sample path in Fig. 1.9.

Fig. 1.11. Sample path of gfBm with scaling function $H_6(t)$ ($n = 4096$).

Fig. 1.12. Estimate of $H_6(t)$ for gfBm sample path in Fig. 1.11.

Fig. 2.1. Partial sums for sample path of gfin with scaling function $H_1(t)$ ($n = 4096$).

Fig. 2.2. Estimate of $H_1(t)$ for gfin sample path in Fig. 2.1.

Fig. 2.3. Partial sums for sample path of gfin with scaling function $H_2(t)$ ($n = 4096$).

Fig. 2.4. Estimate of $H_2(t)$ for gfin sample path in Fig. 2.3.

Fig. 2.5. Partial sums for sample path of gfin with scaling function $H_3(t)$ ($n = 4096$).

Fig. 2.6. Estimate of $H_3(t)$ for gfin sample path in Fig. 2.5.

Fig. 2.7. Partial sums for sample path of gfin with scaling function $H_4(t)$ ($n = 4096$).

Fig. 2.8. Estimate of $H_4(t)$ for gfin sample path in Fig. 2.7.

Fig. 2.9. Partial sums for sample path of gfin with scaling function $H_5(t)$ ($n = 4096$).

Fig. 2.10. Estimate of $H_5(t)$ for gfin sample path in Fig. 2.9.

Fig. 2.11. Partial sums for sample path of gfin with scaling function $H_6(t)$ ($n = 4096$).

Fig. 2.12. Estimate of $H_6(t)$ for gfin sample path in Fig. 2.11.

Fig. 1.1

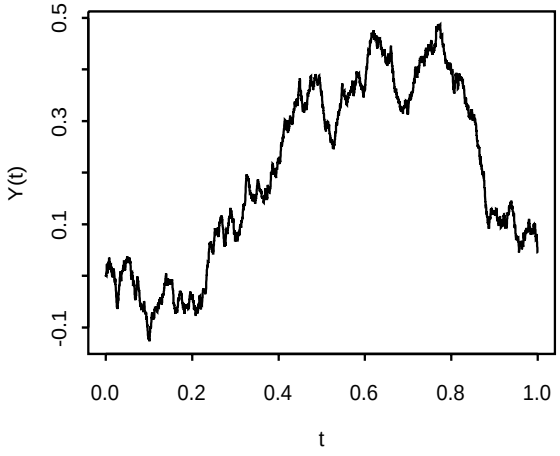


Fig. 1.2

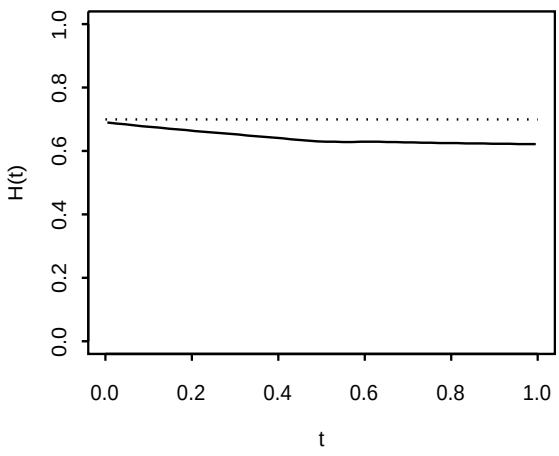


Fig. 1.3

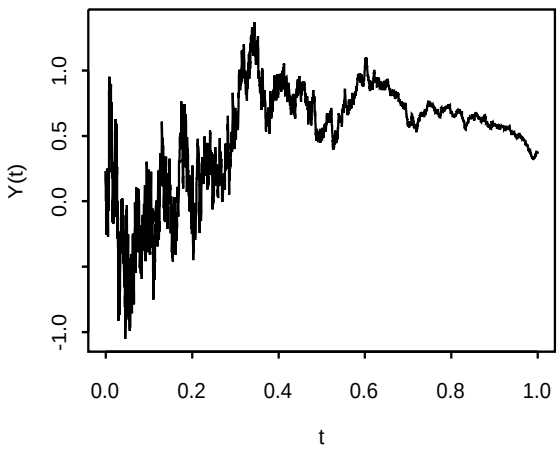


Fig. 1.4

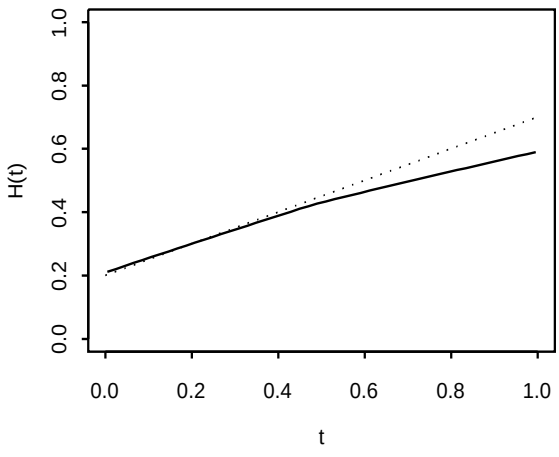


Fig. 1.5

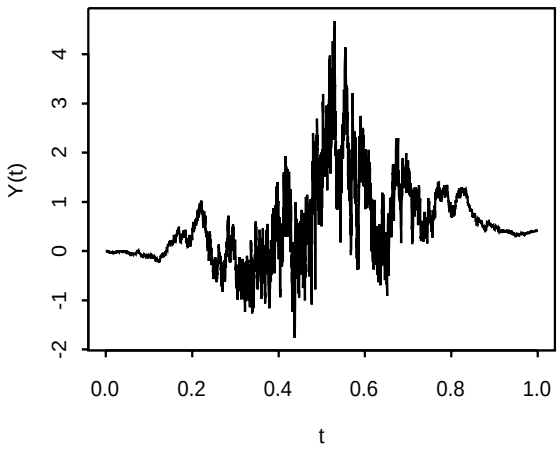


Fig. 1.6

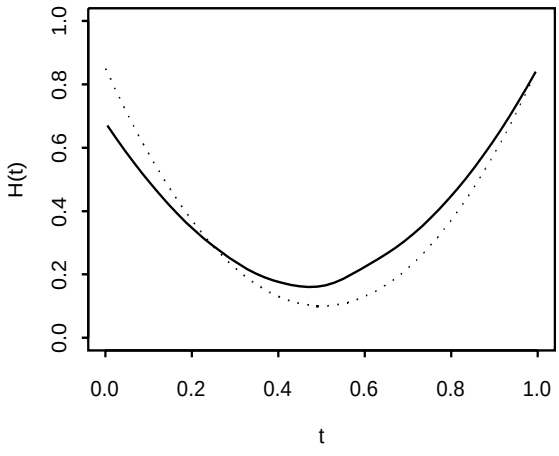


Fig. 1.7

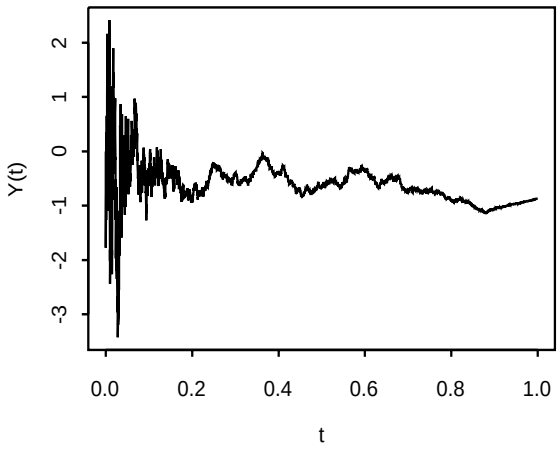


Fig. 1.8

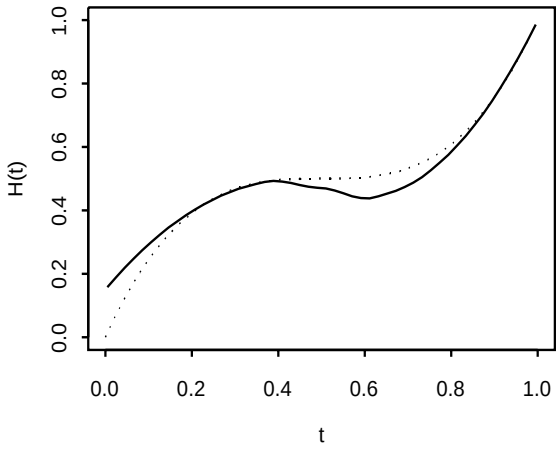


Fig. 1.9

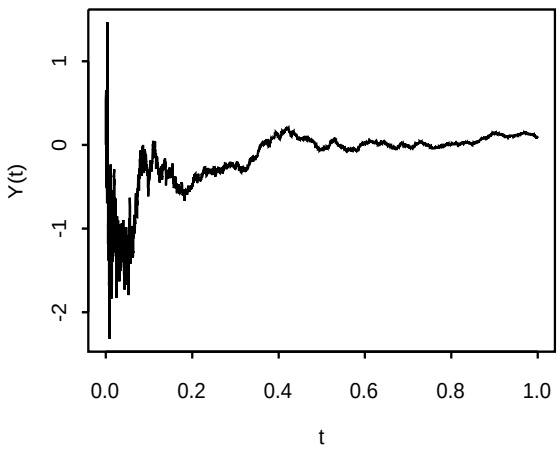


Fig. 1.10

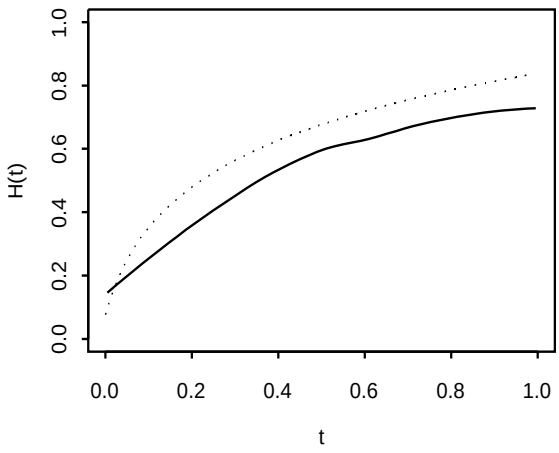


Fig. 1.11

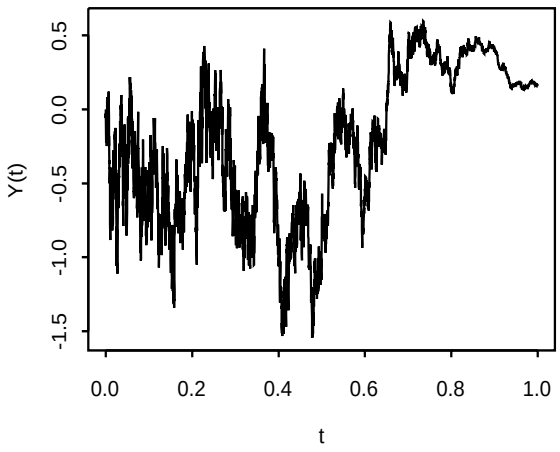


Fig. 1.12

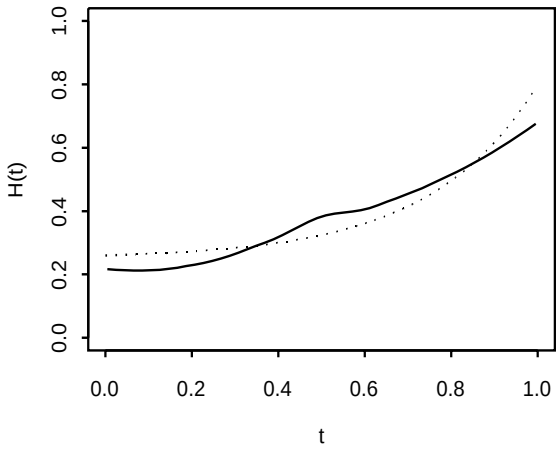


Fig. 2.1

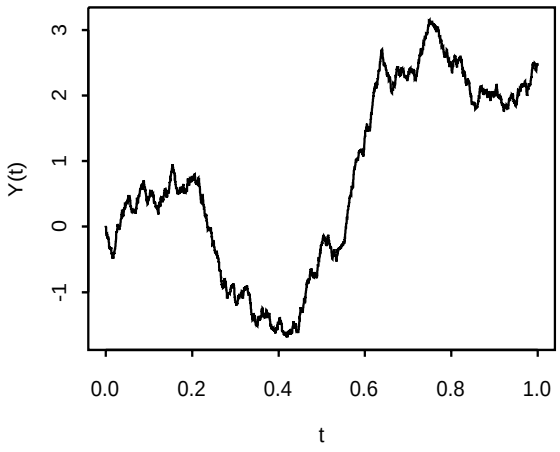


Fig. 2.2

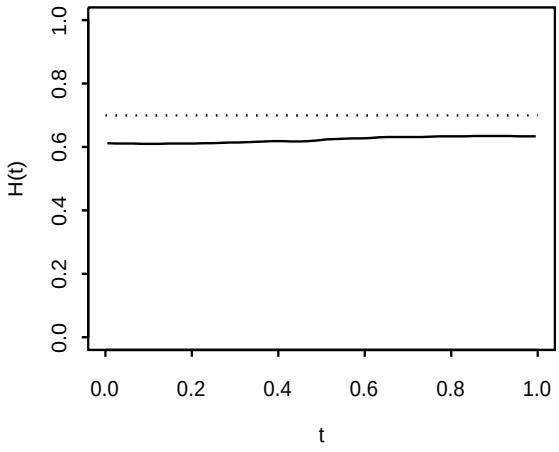


Fig. 2.3

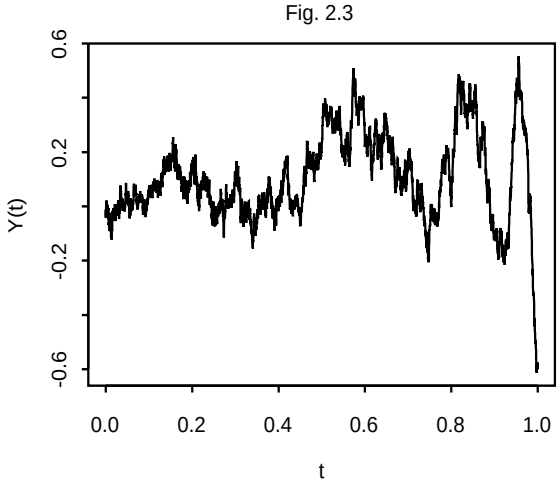


Fig. 2.4

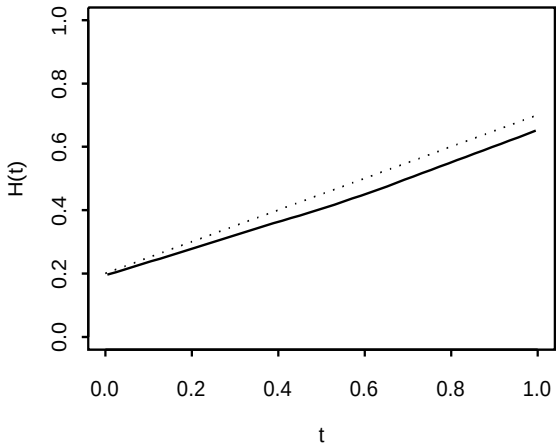


Fig. 2.5

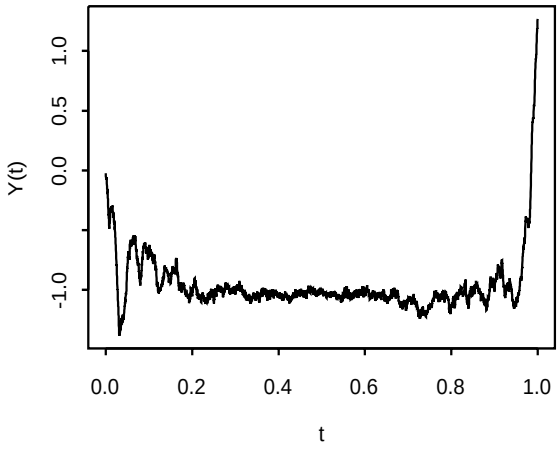


Fig. 2.6

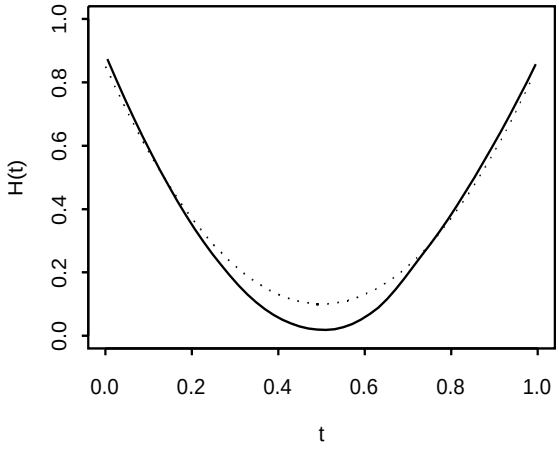


Fig. 2.7

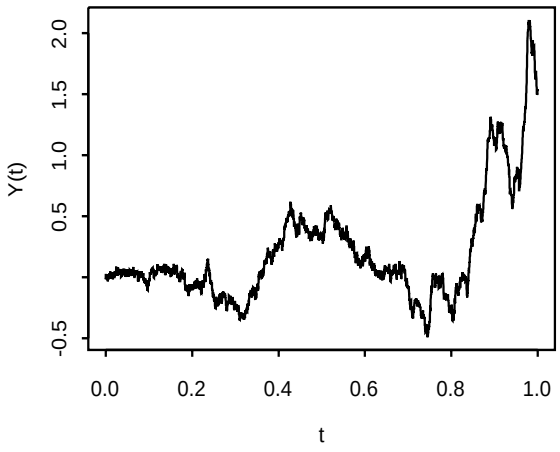


Fig. 2.8

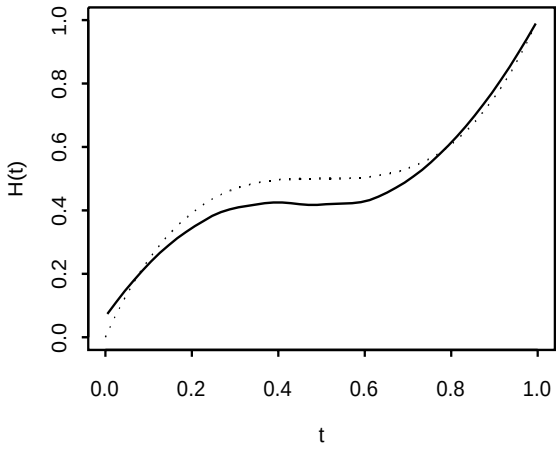


Fig. 2.9

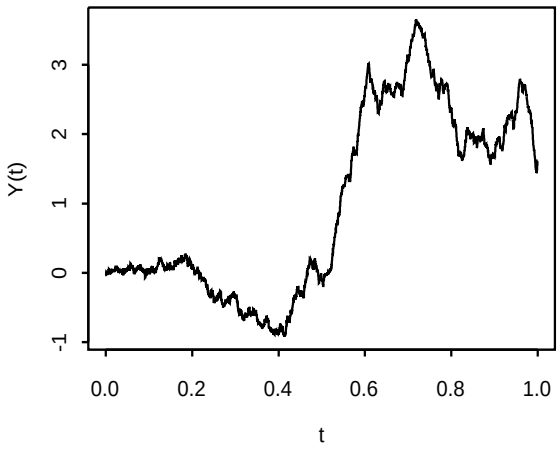


Fig. 2.10

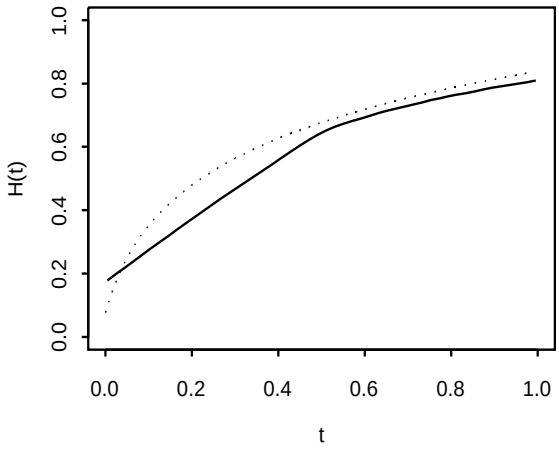


Fig. 2.11

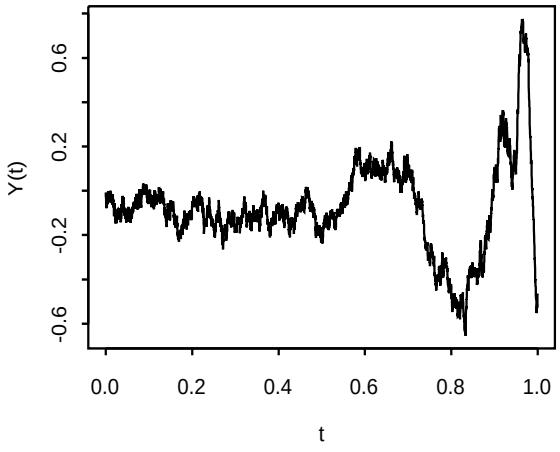


Fig. 2.12

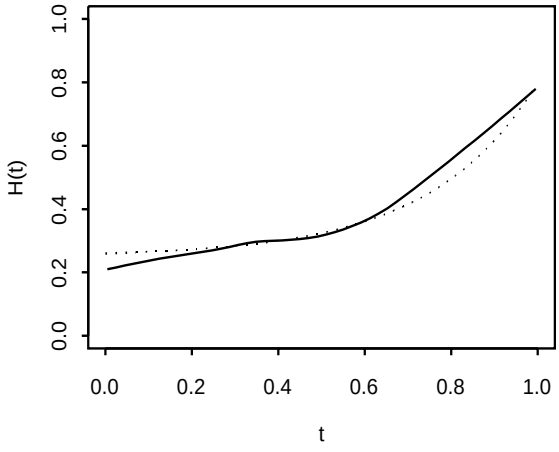


Fig. 3.1

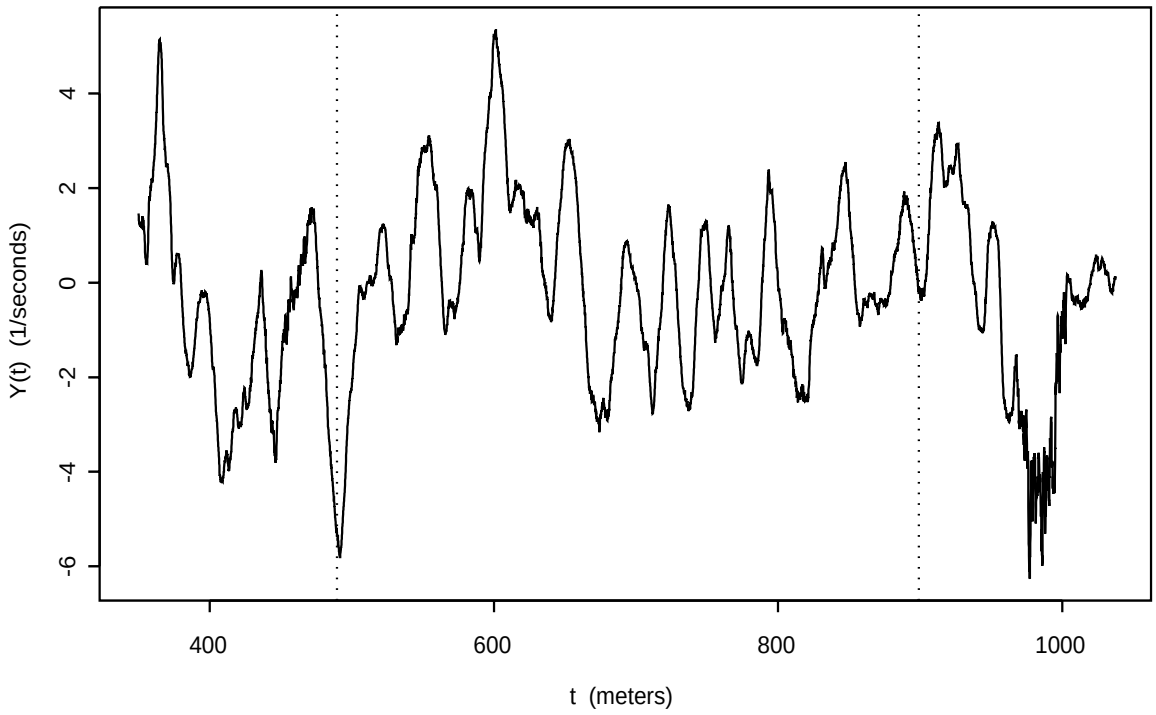

Y(t) (1/seconds)
4
2
0
-2
-4
-6
400
600
800
1000
t (meters)

Fig. 3.2

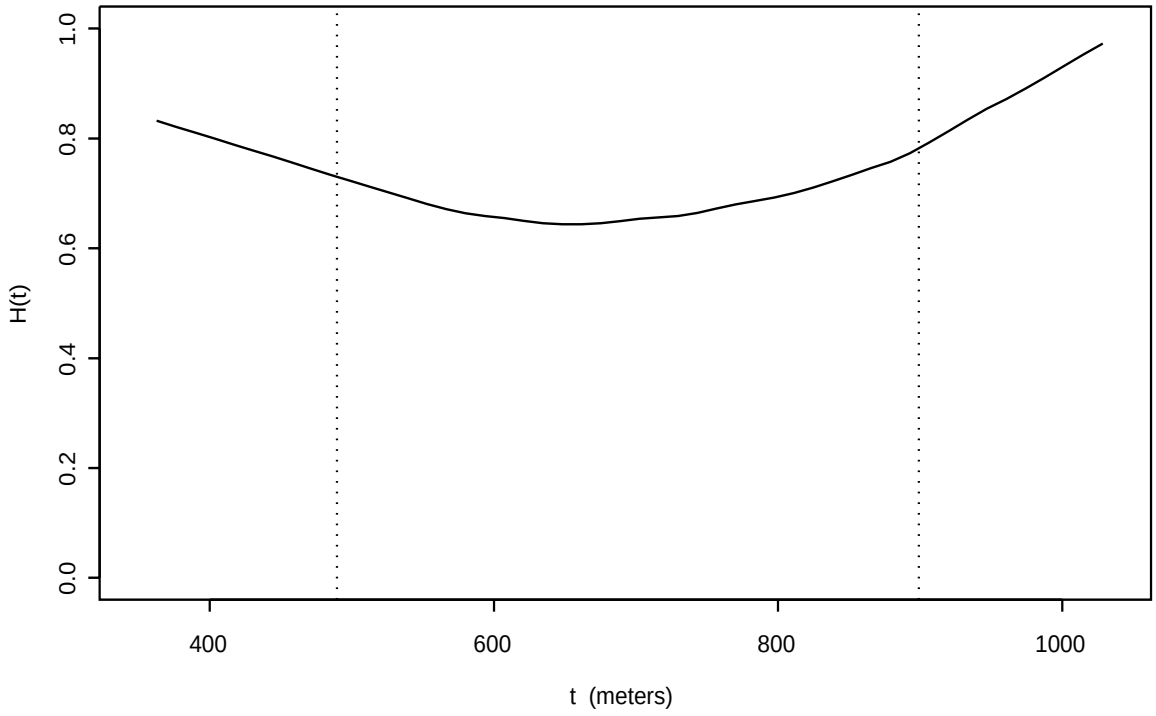

H(t)
1.0
0.8
0.6
0.4
0.2
0.0
400
600
800
1000
t (meters)

Fig. 4.1

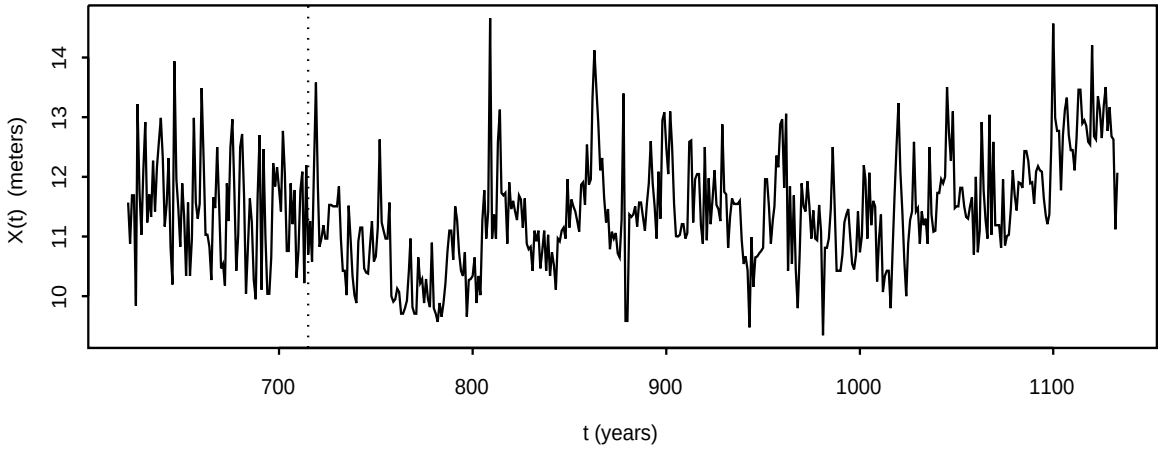


Fig. 4.2

Fig. 4.3

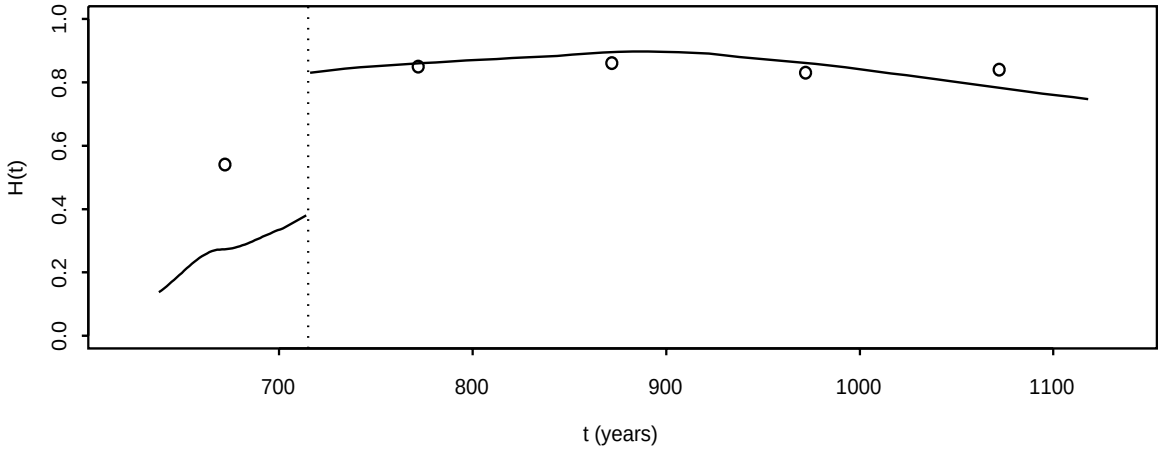


Fig. 5.1

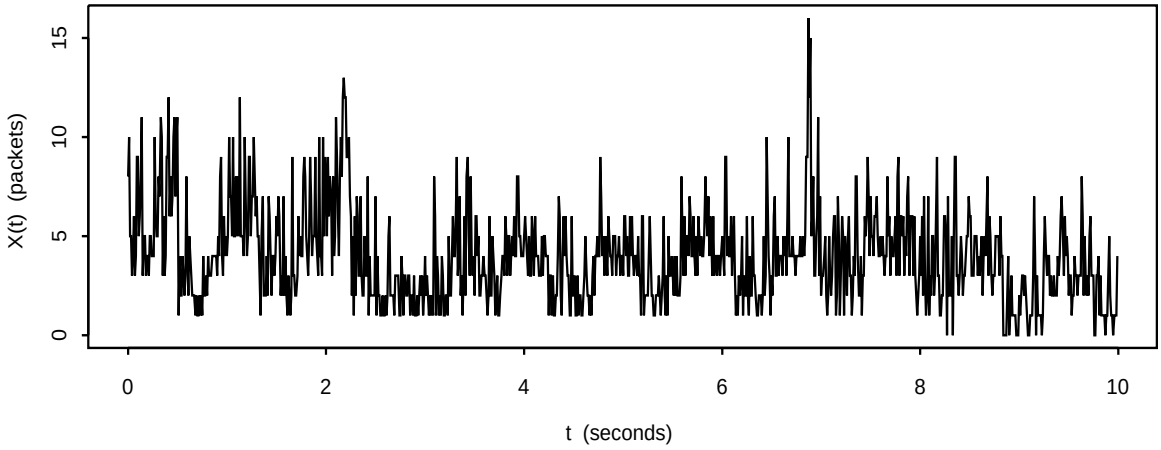


Fig. 5.2

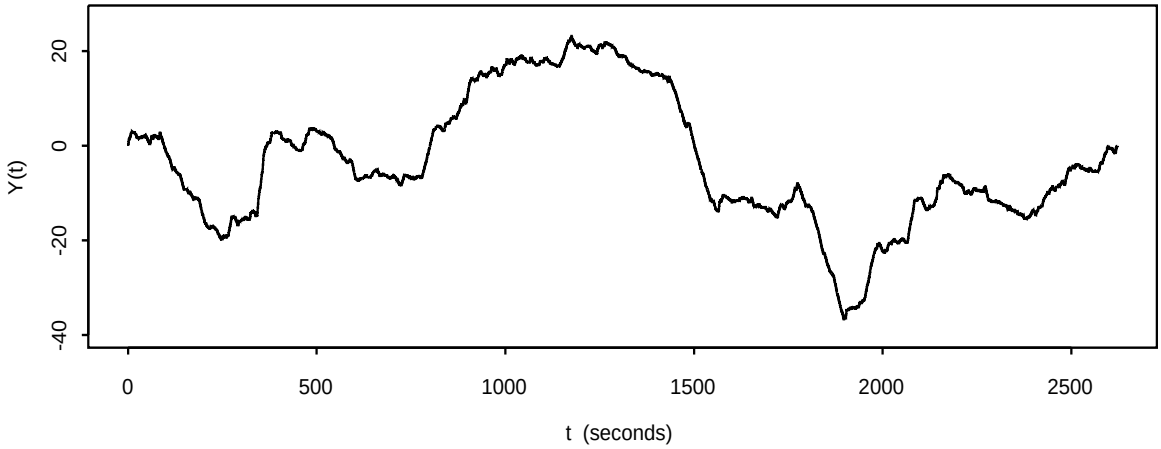
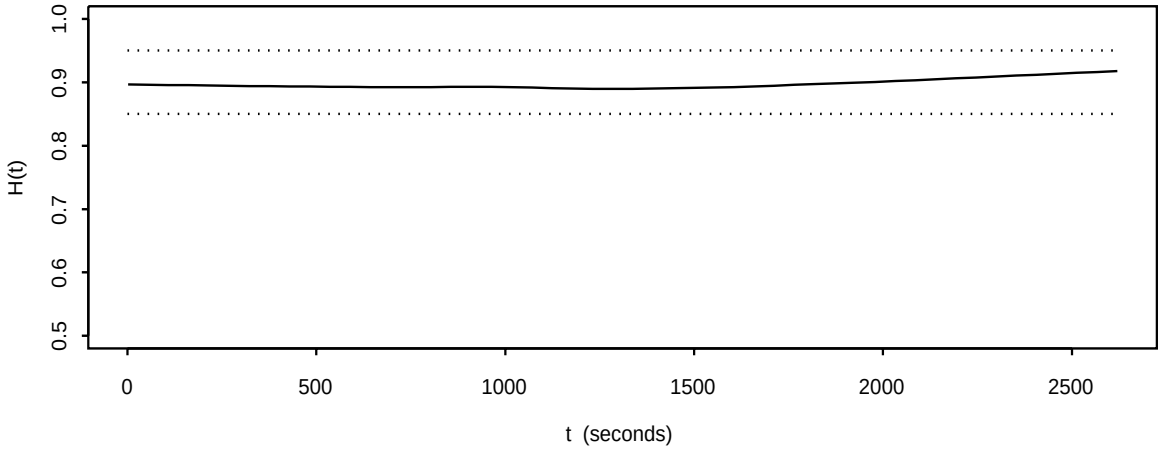


Fig. 5.3



Appendix

In what follows, we present proofs for Theorems 1, 2, 3, and 4 (from section 3), and for Theorem 6 (from section 4). In our proofs, we use K to denote constants that may vary in specification from one usage to the next. In the proof of Theorem 6, these constants may depend on t but not on the scales a.

PROOF OF THEOREM 1.

Assume $s \leq t$. Direct calculations establish

$$\begin{aligned}
R(s,t) &= \int_{-\infty}^{0} \left\{(t-u)^{H(t)-1/2} - (-u)^{H(t)-1/2}\right\} \left\{(s-u)^{H(s)-1/2} - (-u)^{H(s)-1/2}\right\} du \\
&\quad + \int_{0}^{s} (t-u)^{H(t)-1/2} (s-u)^{H(s)-1/2} \, du \\
&= \int_{0}^{\infty} \left\{(t+u)^{H(t)-1/2} - u^{H(t)-1/2}\right\} \left\{(s+u)^{H(s)-1/2} - u^{H(s)-1/2}\right\} du \\
&\quad + \int_{-s}^{0} (t+u)^{H(t)-1/2} (s+u)^{H(s)-1/2} \, du \\
&= \lim_{M \to \infty} A_M,
\end{aligned}$$

where

$$\begin{aligned}
A_M &= \int_{0}^{M} \left\{(t+u)^{H(t)-1/2} - u^{H(t)-1/2}\right\} \left\{(s+u)^{H(s)-1/2} - u^{H(s)-1/2}\right\} du \\
&\quad + \int_{-s}^{0} (t+u)^{H(t)-1/2} (s+u)^{H(s)-1/2} \, du \\
&= \int_{-s}^{M} (t+u)^{H(t)-1/2} (s+u)^{H(s)-1/2} \, du - \int_{0}^{M} (t+u)^{H(t)-1/2} u^{H(s)-1/2} \, du \\
&\quad - \int_{0}^{M} u^{H(t)-1/2} (s+u)^{H(s)-1/2} \, du + \int_{0}^{M} u^{H(s)+H(t)-1} \, du \\
&= \int_{0}^{s+M} (t-s+u)^{H(t)-1/2} u^{H(s)-1/2} \, du - \int_{0}^{M} (t+u)^{H(t)-1/2} u^{H(s)-1/2} \, du \\
&\quad - \int_{0}^{M} u^{H(t)-1/2} (s+u)^{H(s)-1/2} \, du + M^{H(s)+H(t)} / \{H(s)+H(t)\}.
\end{aligned}$$

To make the integrals in A_M converge as $M \to \infty$, we must first subtract $u^{H(s)+H(t)-1}$ and $(1+u)^{H(s)+H(t)-1}$ from each of the integrands, and then perform a change of variables in the integrals. We obtain

$$
\begin{aligned}
A_M &= \int_0^{s+M} \left\{ (t-s+u)^{H(t)-1/2}\, u^{H(s)-1/2} - u^{H(s)+H(t)-1}/2 - (t-s+u)^{H(s)+H(t)-1}/2 \right\} du \\
&\quad - \int_0^{M} \left\{ (t+u)^{H(t)-1/2}\, u^{H(s)-1/2} - u^{H(s)+H(t)-1}/2 - (t+u)^{H(s)+H(t)-1}/2 \right\} du \\
&\quad - \int_0^{M} \left\{ u^{H(t)-1/2}\, (s+u)^{H(s)-1/2} - u^{H(s)+H(t)-1}/2 - (s+u)^{H(s)+H(t)-1}/2 \right\} du \\
&\quad + \left\{ t^{H(s)+H(t)} + s^{H(s)+H(t)} - (t-s)^{H(s)+H(t)} \right\} / \{ 2\,H(s) + 2\,H(t) \} \\
&= -\left\{ (t-s)^{H(s)+H(t)} \int_0^{(s+M)/(t-s)} f_H(u;s,t)\, du \right\} /2 \\
&\quad + \left\{ t^{H(s)+H(t)} \int_0^{M/t} f_H(u;s,t)\, du \right\} /2 + \left\{ s^{H(s)+H(t)} \int_0^{M/s} f_H(u;t,s)\, du \right\} /2 \\
&\quad + \left\{ t^{H(s)+H(t)} + s^{H(s)+H(t)} - (t-s)^{H(s)+H(t)} \right\} / \{ 2\,H(s) + 2\,H(t) \}.
\end{aligned}
$$

The theorem is then established by letting $M \to \infty$. □

Proof of Theorem 2.

(1) We consider $\delta \downarrow 0$ only. By Corollary 1, we have

$$
\begin{aligned}
& V(t_* + \delta, t_*) = \\
& \delta^{H(t_*)+H(t_*+\delta)} \left[\{ H(t_*) + H(t_*+\delta) \}^{-1} + \int_0^\infty f_H(u; t_*, t_*+\delta)\, du \right] \\
& + (t_*+\delta)^{2\,H(t_*+\delta)} \left[\{ 2\,H(t_*+\delta) \}^{-1} + \int_0^\infty f_H(u; t_*+\delta, t_*+\delta)\, du \right] \\
& - (t_*+\delta)^{H(t_*)+H(t_*+\delta)} \left[\{ H(t_*) + H(t_*+\delta) \}^{-1} + \int_0^\infty f_H(u; t_*, t_*+\delta)\, du \right] \\
& + t_*^{2\,H(t_*)} \left[\{ 2\,H(t_*) \}^{-1} + \int_0^\infty f_H(u; t_*, t_*)\, du \right] \\
& - t_*^{H(t_*)+H(t_*+\delta)} \left[\{ H(t_*) + H(t_*+\delta) \}^{-1} + \int_0^\infty f_H(u; t_*+\delta, t_*)\, du \right].
\end{aligned}
$$

Note that $\int_0^\infty f_H(u;s,t)\, du$ is finite and has continuous derivatives with respect to s and t. Also, we can argue that

$$
x^{H(t+\delta)} - x^{H(t)} = \delta\, H'(t)\, x^{H(t)} \log x + \delta^2\, x^{H(t^*)} \left[\{ H'(t^*) \log x \}^2 + H''(t^*) \log x \right],
$$

where t^* is between t_* and $t_* + \delta$, and that

$$
H'(t)\, x^{H(t)} \log x, \qquad x^{H(t^*)} \left[\{ H'(t^*) \log x \}^2 + H''(t^*) \log x \right]
$$

are both bounded. Therefore, we have

$$
\begin{aligned}
&V(t_*+\delta,t_*) = \\
&\delta^{2H(t_*)}\left[\{2H(t_*)\}^{-1}+\int_0^\infty f_H(u;t_*,t_*)\,du\right]+O(\delta^2) \\
&+t_*^{2H(t_*)}\int_0^\infty \{f_H(u;t_*+\delta,t_*+\delta)-f_H(u;t_*,t_*+\delta)+f_H(u;t_*,t_*)-f_H(u;t_*+\delta,t_*)\}\,du \\
&+t_*^{2H(t_*)}\left[\{2H(t_*+\delta)\}^{-1}-2\{H(t_*)+H(t_*+\delta)\}^{-1}+\{2H(t_*)\}^{-1}\right] \\
&+\delta H'(t_*)\,t_*^{2H(t_*)}\log t_*\int_0^\infty \{2f_H(u;t_*+\delta,t_*+\delta)-f_H(u;t_*,t_*+\delta)-f_H(u;t_*+\delta,t_*)\}\,du \\
&+\delta H'(t_*)\,t_*^{2H(t_*)}\log t_*\,[2\{2H(t_*+\delta)\}^{-1}-2\{H(t_*)+H(t_*+\delta)\}^{-1}] \\
&+2\delta H(t_*)\,t_*^{2H(t_*)-1}\int_0^\infty \{f_H(u;t_*+\delta,t_*+\delta)-f_H(u;t_*,t_*+\delta)\}\,du \\
&+2\delta H(t_*)\,t_*^{2H(t_*)-1}[\{2H(t_*+\delta)\}^{-1}-\{H(t_*)+H(t_*+\delta)\}^{-1}] \\
&=\delta^{2H(t_*)}\left[\{2H(t_*)\}^{-1}+\int_0^\infty f_H(u;t_*,t_*)\,du\right]+O(\delta^2).
\end{aligned}
$$

(2) By the arguments in the proof of Theorem 1, we can establish that

$$
\begin{aligned}
E\{Y(t_*+\delta)\tilde{Y}(t_*+\delta)\} &= (t_*+\delta)^{H(t_*)+H(t_*+\delta)}\Big[\{H(t_*)+H(t_*+\delta)\}^{-1} \\
&+\int_0^\infty \{f_H(u;t_*,t_*+\delta)+f_H(u;t_*+\delta,t_*)\}/2\,du\Big].
\end{aligned}
$$

Using Corollary 1 and the arguments from part (1), we can then prove $\mathrm{Var}\{Y(t_*+\delta)-\tilde{Y}(t_*+\delta)\}$ is of order δ^2. □

PROOF OF THEOREM 3.

Since $\tilde{Y}(t)$ is fBm with self-similarity parameter $H(t_*)$, by the law of the iterated logarithm for fBm (Oodaira, 1972; Csáki, 1980; Goodman and Kuelbs, 1991; Monrad and Rootzén, 1995), there exist constants c and C depending only on $H(t_*)$ such that

$$
\liminf_{\delta\to 0}\sup\left\{\frac{|\tilde{Y}(s)-\tilde{Y}(t)|}{(\delta/\log|\log\delta|)^{H(t_*)}};s,t\in(t_*-\delta,t_*+\delta)\right\}=c,
$$

$$
\limsup_{\delta\to 0}\sup\left\{\frac{|\tilde{Y}(s)-\tilde{Y}(t)|}{\delta^{H(t_*)}/(\log|\log\delta|)^{1/2}};s,t\in(t_*-\delta,t_*+\delta)\right\}=C.
$$

Let $Z(t)=Y(t)-\tilde{Y}(t)$, and

$$
m(t_*,\delta)=\sup\{|Z(s)-Z(t)|;s,t\in(t_*-\delta,t_*+\delta)\}.
$$

Then it is sufficient to prove

$$\lim_{\delta \to 0} \frac{m(t_*, \delta)}{\delta^{H(t_*)} / \log \, |\log \, \delta|} = 0. \tag{A.1}$$

By Theorem 2.1 of Monrad and Rootźen (1995) and by Theorem 2, for $2\,H(t_*) < \gamma < 2$, we have that

$$P\{m(t_*, \delta) > \varepsilon\} \leq 1 - \exp\left(-K\,\delta\,\varepsilon^{-2/\gamma}\right).$$

Let $0 < \zeta < \gamma/2 - H(t_*)$, and for $k = 1, 2, \cdots$, set

$$\delta_k = k^{1/[\{H(t_*)+\zeta\}/\gamma - 1/2]} \qquad \text{and} \qquad \varepsilon_k = \delta_k^{H(t_*)+\zeta}.$$

Then,

$$\begin{aligned} \sum_k P\left\{m(t_*, \delta_k) > \delta_k^{H(t_*)+\zeta}\right\} &\leq \sum_k \left(1 - \exp\left[-K\,\delta_k^{1-2\,\{H(t_*)+\zeta\}/\gamma}\right]\right) \\ &\sim \sum_k K\,\delta_k^{1-2\,\{H(t_*)+\zeta\}/\gamma} \\ &= K \sum_k k^{-2} < \infty. \end{aligned}$$

By the Borel-Cantelli lemma, it follows that $m(t_*, \delta_k) \leq \delta_k^{H(t_*)+\zeta}$. Furthermore, for $\delta_{k+1} < \delta \leq t_k$,

$$m(t_*, \delta) \leq m(t_*, \delta_k) \leq \delta_k^{H(t_*)+\zeta} \leq \delta^{H(t_*)+\zeta}\,(\delta_k/\delta_{k+1})^{H(t_*)+\zeta}.$$

Thus, for small δ, $m(t_*, \delta) \leq 2\,\delta^{H(t_*)+\zeta}$. This establishes (A.1). □

Proof of Theorem 4.

Since $D_{[a,b]}(Y) \leq \Delta_{[a,b]}(Y)$, we need to show $D_{[a,b]}(Y) \geq 2 - H_{ab}$ and $\Delta_{[a,b]}(Y) \leq 2 - H_{ab}$.

(1) $\Delta_{[a,b]}(Y) \leq 2 - H_{ab}$. It is sufficient to show that for any $\gamma < H_{ab}$, we have $\Delta_{[a,b]}(Y) \leq 2 - \gamma$.

Suppose that for any $\gamma < H_{ab}$,

$$P\left[\sup\left\{|Y(t) - Y(s)|/|s - t|^\gamma; s, t \in [a, b]\right\} < \infty\right] = 1. \tag{A.2}$$

We will then have $\Delta_{[a,b]}(Y) \leq 2 - \gamma$ (Adler, 1981, Chapter 8; Tricot, 1995, Chapter 12). Now for $t - s$ bounded below from zero, with probability one, $|Y(t) - Y(s)|/|s - t|^\gamma < \infty$, so we need only show (A.2) for small $t - s$.

Theorem 2 implies that there exists a positive constant K such that for $s, t \in [a, b]$ and $t - s \to 0$,

$$V(t, s) \leq K \, |t - s|^{2\, H_{ab}}.$$

Let

$$m(\delta) = \sup \{|Y(t) - Y(s)|; s, t \in [a, b], |t - s| \leq \delta\}.$$

By Theorem 2.1 of Monrad and Rootźen (1995), we have

$$P\{m(\delta) \geq \varepsilon\} \leq 1 - \exp\left(-K \, \delta \, \varepsilon^{-1/H_{ab}}\right).$$

For $k = 1, 2, \cdots,$ set $\delta_k = k^{2/(\gamma/H_{ab}-1)}$ and $\varepsilon_k = \delta_k^{\gamma}$. Then

$$\begin{aligned} \sum_k P\{m(\delta_k) \geq \delta_k^{\gamma}\} &\leq \sum_k \left\{1 - \exp\left(-K \, \delta_k \, \varepsilon_k^{-1/H_{ab}}\right)\right\} \\ &\sim \sum_k K \, \delta_k^{1-\gamma/H_{ab}} \\ &= K \sum_k k^{-2} < \infty. \end{aligned}$$

By the Borel-Cantelli lemma, it follows that $m(\delta_k) \, \delta_k^{\gamma} \leq 1$. Furthermore, for $\delta_{k+1} < \delta \leq t_k$,

$$\frac{m(\delta)}{\delta^{\gamma}} \leq \frac{m(\delta_k)}{\delta_k^{\gamma}} \left(\frac{\delta_k}{\delta_{k+1}}\right)^{\gamma}.$$

Since $\delta_k / \delta_{k+1} \to 1$, we have $m(\delta)/\delta^{\gamma} \leq 2$. Finally,

$$|Y(t) - Y(s)|/|t - s|^{\gamma} \leq m(|t - s|)/|t - s|^{\gamma}.$$

This establishes (A.2).

(2) $D_{[a,b]}(Y) \geq 2 - H_{ab}$. It it sufficient to show that for any $\gamma > H_{ab}$, we have $D_{[a,b]}(Y) \geq 2 - \gamma$.

Suppose $t_* \in [a, b]$ is such that $H_{ab} = H(t_*)$. There exists an $\epsilon > 0$ such that for $t \in [t_* - \epsilon, t_* + \epsilon]$, we have $\gamma > H(t) + \epsilon$. Since $D_{[a,b]}(Y) \geq D_{[t_*-\epsilon, t_*+\epsilon]}(Y)$, to prove $D_{[a,b]}(Y) \geq 2 - \gamma$, it suffices to show that $D_{[t_*-\epsilon, t_*+\epsilon]}(Y) \geq 2 - \gamma$, which is implied by

$$\int_{t_*-\epsilon}^{t_*+\epsilon} \int_{t_*-\epsilon}^{t_*+\epsilon} E\left\{|Y(t) - Y(s)|^2 + |s - t|^2\right\}^{\gamma/2-1} \, ds \, dt < \infty.$$

(See Adler, 1981, Chapter 8, Lemma 8.2.4.)

Note that $Y(t) - Y(s)$ follows a normal distribution with mean zero. If $\phi(y)$ denotes the probability density function of the standard normal distribution, we then have

$$\begin{aligned}
&\int_{t_*-\epsilon}^{t_*+\epsilon} \int_{t_*-\epsilon}^{t_*+\epsilon} E\left\{|Y(t) - Y(s)|^2 + |s-t|^2\right\}^{\gamma/2-1} \, ds\, dt \\
&= \int_{t_*-\epsilon}^{t_*+\epsilon} \int_{t_*-\epsilon}^{t_*+\epsilon} \int \left\{y^2\, V(s,t) + |s-t|^2\right\}^{\gamma/2-1} \phi(y) dy\, ds\, dt \\
&\le K \int_{t_*-\epsilon}^{t_*+\epsilon} \int_{-2\,\epsilon}^{2\,\epsilon} \int \left\{y^2\, V(t+u,t) + u^2\right\}^{\gamma/2-1} dy\, du\, dt.
\end{aligned}$$

Now the integral over $[-2\,\epsilon, 2\,\epsilon]$ with respect to u need only be considered over $[-\delta, \delta]$ for small $0 < \delta \le 2\,\epsilon$. By Theorem 2, we have that for small u, $V(t+u,t) \sim C(t)\,|u|^{2\,H(t)}$, with $C(t)$ bounded below from zero and above from infinity. Thus it is sufficient to establish the finiteness of

$$\int_{t_*-\epsilon}^{t_*+\epsilon} \int_{-\delta}^{\delta} \int \left\{y^2\, |u|^{2\,H(t)} + u^2\right\}^{\gamma/2-1} dy\, du\, dt.$$

This integral in turn is bounded by

$$\begin{aligned}
&8\,\delta \int_{t_*-\epsilon}^{t_*+\epsilon} \int_0^{\delta} \int_0^{\infty} \left\{y^2\, u^{2\,H(t)} + u^2\right\}^{\gamma/2-1} dy\, du\, dt \\
&= 8\,\delta \int_{t_*-\epsilon}^{t_*+\epsilon} \int_0^{\delta} u^{H(t)\,(\gamma-2)} \int_0^{\infty} \left\{y^2 + u^{2-2\,H(t)}\right\}^{\gamma/2-1} dy\, du\, dt \\
&\le K \int_{t_*-\epsilon}^{t_*+\epsilon} \int_0^{\delta} u^{H(t)\,(\gamma-2)}\, u^{\{1-H(t)\}\,(\gamma-1)}\, du\, dt \\
&= K \int_{t_*-\epsilon}^{t_*+\epsilon} \int_0^{\delta} u^{\gamma-H(t)-1}\, du\, dt \\
&\le K \int_{t_*-\epsilon}^{t_*+\epsilon} \int_0^{\delta} u^{\epsilon-1}\, du\, dt \\
&= K\,\epsilon\,\delta^{\epsilon} < \infty. \qquad \square
\end{aligned}$$

PROOF OF THEOREM 6.

The theorem can be easily established by proving

$$E\left\{\hat{H}(t)\right\} \to H(t) \qquad \text{and} \qquad \mathrm{Var}\left\{\hat{H}(t)\right\} \to 0. \tag{A.3}$$

The result then follows by Chebyshev's inequality.

Recall from section 4 that the bivariate data values which lead to $\hat{H}(t)$ are denoted by $x_j = \log(2^{-j})$ and $y_j = y_t(2^{-j})$ for $j = 1, \ldots, k$, where $y_t(a) = \log\{|TY(a,t)|^2\}$. Thus,

$\hat{H}(t)$ is based on a bivariate sample of size k. We will prove (A.3) by demonstrating that as $k \to \infty$,

$$E\{\hat{H}(t)\} = H(t) + o(k^{-1}), \qquad \mathrm{Var}\{\hat{H}(t)\} = O(k^{-3}). \tag{A.4}$$

To ease notation, we suppress the time index t in the specifications of $y_t(a)$, $H(t)$, and $\hat{H}(t)$, and write these objects simply as $y(a)$, H, and $\hat{H}$, respectively.

Direct computations show that

$$\sum_{j=1}^{k}(x_j - \bar{x})^2 = (\log\ 2)^2\ \left\{\sum_{j=1}^{k} j^2 - k\,(k+1)^2/4\right\} = (\log\ 2)^2\ k\,(k+1)\,(k-1)/12 \sim k^3, \tag{A.5}$$

and

$$\sum_{j=1}^{k} |x_j - \bar{x}| = (\log\ 2)\ \sum_{j=1}^{k} |j - (k+1)/2| \sim k^2. \tag{A.6}$$

By (4.2), we have as $a \to 0$,

$$E\left\{|TY(a,t)|^2\right\} = C_1\, a^{2\,H+1}\ \{1 + o(1)\},$$

and thus

$$\log\left[E\left\{|TY(a,t)|^2\right\}\right] = \log C_1 + (2\,H+1)\ \log\ a + o(1).$$

The preceding relation and (4.3) together imply

$$E(y_j) = c + (2\,H+1)\,x_j + o(1), \tag{A.7}$$

where $c = \log C_1 + C_2$. From (4.5) along with (A.5), (A.6), and (A.7), we obtain

$$\begin{aligned} E(\hat{H}) &= \frac{\sum(x_j - \bar{x})\ E(y_j - \bar{y})}{2\ \sum(x_j - \bar{x})^2} - \frac{1}{2} \\ &= H + o\left(\frac{\sum |x_j - \bar{x}|}{\sum(x_j - \bar{x})^2}\right) \\ &= H + o(k^{-1}). \end{aligned}$$

This establishes the first result of (A.4). We now derive the second result.

Utilizing (4.1), the covariance property (2.3), and the first vanishing moment of the mother wavelet ψ, we have

$$
\begin{aligned}
&\mathrm{Cov}\{TY(a_1,t), TY(a_2,t)\} \\
&= \frac{1}{(a_1\,a_2)^{1/2}} \int\int \psi\left(\frac{u-t}{a_1}\right)\,\psi\left(\frac{v-t}{a_2}\right)\,E\{Y(u)\,Y(v)\}\,du\,dv \\
&= (a_1\,a_2)^{1/2} \int\int \psi(x)\,\psi(y)\,E\{Y(t+a_1\,x)\,Y(t+a_2\,y)\}\,dx\,dy \\
&= (a_1\,a_2)^{1/2} \int\int \psi(x)\,\psi(y)\,\Sigma_t(a_1\,x, a_2\,y)\,dx\,dy \\
&\sim (a_1\,a_2)^{1/2} \int\int \psi(x)\,\psi(y)\,\left\{\Sigma_t(0,0) - C(t)\,|a_1\,x - a_2\,y|^{2\,H}\right\}\,dx\,dy \\
&= -C(t)\,(a_1\,a_2)^{1/2} \int\int \psi(x)\,\psi(y)\,|a_1\,x - a_2\,y|^{2\,H}\,dx\,dy \\
&= -C(t)\,a_1^{2\,H+1/2}\,a_2^{1/2} \int\int \psi(u + y\,a_2/a_1)\,\psi(y)\,|u|^{2\,H}\,du\,dy, \quad \text{as } a_1, a_2 \to 0. \qquad \text{(A.8)}
\end{aligned}
$$

Now assume the mother wavelet ψ has M vanishing moments and is M times continuously differentiable where $M \geq 2$. We take an M^{th}-order expansion of $\psi(u + y\,a_2/a_1)$ about u to establish that as $(a_2/a_1) \to 0$,

$$
\psi(u + y\,a_2/a_1) = \psi(u) + y\psi'(u)(a_2/a_1) + \ldots + (1/M!)y^M\psi^{(M)}(u)(a_2/a_1)^M\{1 + o(1)\}.
$$

This expansion allows us to show

$$
\int\int \psi(u + y\,a_2/a_1)\,\psi(y)\,|u|^{2\,H}\,du\,dy \sim K(a_2/a_1)^M, \qquad \text{as } (a_2/a_1) \to 0.
$$

Using the preceding in (A.8), we see that as $a_1, a_2 \to 0$ with $(a_2/a_1) \to 0$,

$$
\mathrm{Cov}\{TY(a_1,t), TY(a_2,t)\} \sim K\,a_1^{2\,H+1/2}\,a_2^{1/2}\,(a_2/a_1)^M.
$$

This result implies that for small a_1, a_2 chosen such that $a_2 < a_1$,

$$
|\mathrm{Corr}\{TY(a_1,t), TY(a_2,t)\}| \leq K\,(a_2/a_1)^{M-H}. \qquad \text{(A.9)}
$$

Now for a bivariate normal pair of random variables (W, Z) with $E(W) = E(Z) = 0$, it can be shown that

$$
|\mathrm{Corr}\{\log(W^2), \log(Z^2)\}| \leq [\mathrm{Corr}(W, Z)]^2\,.
$$

It therefore follows from (A.9) that for small $a_2 < a_1$,

$$
|\mathrm{Corr}\{y(a_1), y(a_2)\}| \;\leq\; [\mathrm{Corr}\{TY(a_1,t), TY(a_2,t)\}]^2 \;\leq\; K\,(a_2/a_1)^{2\,M-2\,H}. \qquad \text{(A.10)}
$$

Furthermore, if W is normal with variance σ^2, the variance of $\log(W^2)$ does not depend on σ^2, as can be seen by writing $\log(W^2) = \log\{(W/\sigma)^2\} + \log\sigma^2$. Utilizing this fact along with (A.10), we have that for small $a_2 < a_1$,

$$|\mathrm{Cov}\{y(a_1), y(a_2)\}| \leq K\,(a_2/a_1)^{2\,M-2\,H}.$$

From the preceding, we have

$$|\mathrm{Cov}(y_i, y_j)| \leq K\,2^{|i-j|\,(2\,H-2\,M)}. \tag{A.11}$$

With (A.11), we can write

$$\begin{aligned}
|\sum_{ij} x_i\,x_j\,\mathrm{Cov}(y_i, y_j)| &\leq K\sum_{ij} i\,j\,2^{|i-j|\,(2\,H-2\,M)} \\
&\leq K\sum_{i\leq j} i\,j\,2^{(j-i)\,(2\,H-2\,M)} \\
&= K\sum_{i=1}^{k}\left[i\,2^{-i\,(2\,H-2\,M)}\left\{\sum_{j=i}^{k} j\,2^{j\,(2\,H-2\,M)}\right\}\right] \\
&\leq K\sum_{i=1}^{k}\left[i\,2^{-i\,(2\,H-2\,M)}\left\{i\,2^{i\,(2\,H-2\,M)}\right\}\right] \\
&= K\sum_{i=1}^{k} i^2 \\
&\sim k^3.
\end{aligned} \tag{A.12}$$

Similarly, we can show

$$\bar{x}\,|\sum_{ij} x_i\,\mathrm{Cov}(y_i, y_j)| \sim k^3, \qquad (\bar{x})^2\,|\sum_{ij}\mathrm{Cov}(y_i, y_j)| \sim k^3. \tag{A.13}$$

By (A.12) and (A.13), we have

$$\begin{aligned}
\sum_{ij}(x_i-\bar{x})\,(x_j-\bar{x})\,\mathrm{Cov}(y_i, y_j) &= \sum_{ij} x_i\,x_j\,\mathrm{Cov}(y_i, y_j) - \bar{x}\,\sum_{ij}(x_i+x_j)\,\mathrm{Cov}(y_i, y_j) \\
&\quad + (\bar{x})^2\,\sum_{ij}\mathrm{Cov}(y_i, y_j) \\
&\sim k^3,
\end{aligned}$$

Finally, using the preceding along with (4.5) and (A.5), we have

$$\mathrm{Var}(\hat{H}) = \frac{\sum_{ij}(x_i-\bar{x})\,(x_j-\bar{x})\,\mathrm{Cov}(y_i, y_j)}{4\{\sum(x_i-\bar{x})^2\}^2} \sim k^{-3}. \qquad \square$$

References

Abry, P., P. Flandrin, M.S. Taqqu and D. Veitch (2001). Self-similarity and long-range dependence through the wavelet lens. In: P. Doukhan, G. Oppenheim and M.S. Taqqu, eds., *Long-Range Dependence: Theory and Applications.* Birkhäuser, Boston.

Abry, P., P. Flandrin, M.S. Taqqu and D. Veitch (2000). Wavelets for the analysis, estimation and synthesis of scaling data. In: K. Park and W. Willinger, eds., *Self-Similar Network Traffic and Performance Evaluation.* Wiley, New York.

Abry, P. and F. Sellan (1996). The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation. *Applied and Computational Harmonic Analysis* **3**, 377–383.

Adler, R.L. (1981). *The Geometry of Random Fields.* Wiley, New York.

Bardet, J.-M., G. Lang, E. Moulines and P. Soulier (2000). Wavelet estimator of long-range dependent processes. *Statistical Inference for Stochastic Processes* **3**, 85–99.

Bardet, J.-M., G. Lang, G. Oppenheim, A. Philippe, S. Stoev and M.S. Taqqu (2001). Semi-parametric estimation of the long-range dependence parameter: A survey. In: P. Doukhan, G. Oppenheim and M.S. Taqqu, eds., *Long-Range Dependence: Theory and Applications.* Birkhäuser, Boston.

Barnsley, M.F. (1993). *Fractals Everywhere (Second Edition).* Academic Press, Boston.

Beran, J. (1994). *Statistics for Long Memory Processes.* Chapman and Hall, New York.

Buldyrev, S.V., A.L. Goldberger, S. Havlin, C-K. Peng, H.E. Stanley, M.H.R. Stanley and M. Simons (1993). Fractal landscapes and molecular evolution: Modeling the Moyosin heavy chain gene family. *Biophysical Journal* **65**, 2673–2679.

Chen, G., P. Hall and D.S. Poskitt (1995). Periodogram-based estimators of fractal properties. *The Annals of Statistics* **23**, 1684–1711.

Cohen, A., I. Daubechies, B. Jawerth and P. Vail (1993a). Multiresolution analysis, wavelets and fast algorithms on an interval. *Comptes Rendus des Séances de l'Académie des Sciences, Série I, Mathématique* **316**, 417–421.

Cohen, A., I. Daubechies and P. Vail (1993b). Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis* **1**, 54–81.

Comte, F. (1996). Simulation and estimation of long memory continuous time models. *Journal of Time Series Analysis* **17**, 19–36.

Constantine, A.G. and P. Hall (1994). Characterizing surface smoothness via estimation of effective fractal dimension. *Journal of the Royal Statistical Society B* **56**, 97–113.

Csáki, E. (1980). A relation between Chung's and Strassen's laws of the iterated logarithm. *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete* **54**, 287–301.

Daubechies, I. (1992). *Ten Lectures on Wavelets.* CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.

Daubechies, I. (1994). Two recent results on wavelets: Wavelet bases for the interval, and biorthogonal wavelets diagonalizing the derivative operator. In: L.L. Schumaker and G. Webb, eds., *Recent Advances in Wavelet Analysis.* Academic Press, Boston.

Donoho, D.L. and I.M. Johnstone (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81**, 425–455.

Farge, M., C.R. Hunt and J.C. Vassilicos (1993). *Wavelets, Fractals and Fourier Transformations.* Clarendon Press, Oxford.

Flandrin, P. and P. Gonçalvès (1994). From wavelets to time-scale energy distributions. In: L.L. Schumaker and G. Webb, eds., *Recent Advances in Wavelet Analysis.* Academic Press, Boston.

Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long-memory time series models. *Journal of Time Series Analysis* **4**, 221–237.

Gonçalvès, P. and Flandrin, P. (1993). Bilinear time-scale analysis applied to local scaling exponents estimation. In: Y. Meyer and S. Roques, eds., *Progress in Wavelet Analysis and Applications.* Frontieres, Paris.

Goodman, V. and J. Kuelbs (1991). Rates of clustering for some Gaussian self-similar processes. *Probability Theory and Related Fields* **88**, 47–75.

Hall, P., H.L. Koul and B.A. Turlach (1997). Note on convergence rates of semiparametric estimators of dependence index. *The Annals of Statistics* **25**, 1725–1739.

Hwang, W.L. and S. Mallat (1994). Characterization of self-similar multifractals with wavelet maxima. *Applied and Computational Harmonic Analysis* **4**, 316–328.

Kent, J.T. and A.T.A. Wood (1997). Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *Journal of the Royal Statistical Society B* **59**, 679–700.

Kolmogorov, A.N. (1941). Local structure of turbulence in an incompressible viscous fluid

at very large Reynolds numbers. Translation in: S.K. Friedlander and L. Topper, eds., *Turbulence: Classic Papers on Statistical Theory* (1961). Interscience Publishers Inc., New York.

Leland, W.E., M.S. Taqqu, W. Willinger and D.V. Wilson (1994). On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking* **2**, 1–15.

Mandelbrot, B.B. (1983). *The Fractal Geometry of Nature (Updated and Augmented Edition).* Freeman, New York.

Mandelbrot, B.B. and J.W. van Ness (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review* **10**, 422–437.

Mandelbrot, B.B. and J.R. Wallis (1968). Noah, Joseph and operational hydrology. *Water Resources Research* **4**, 909–918.

Mandelbrot, B.B. and J.R. Wallis (1969). Computer experiments with fractional Gaussian noises. *Water Resources Research* **5**, 228–267.

McCoy, E.J. and A.T. Walden (1996). Wavelet analysis and synthesis of stationary long-memory processes. *Journal of Computational and Graphical Statistics* **5**, 26–56.

Monrad, D. and H. Rootzén (1995). Small values of Gaussian processes and functional laws of the iterated logarithm. *Probability Theory and Related Fields* **101**, 173–192.

Moulines, E. and P. Soulier (1999). Broadband log-periodogram regression of time series with long-range dependence. *The Annals of Statistics* **27**, 1415–1439.

Oodaira, H. (1972). On Strassen's version of the law of the iterated logarithm for Gaussian processes. *Zeitschrift fr̈ Wahrscheinlichkeitstheorie und Verwandte Gebiete* **21**, 289–299.

Ossadnik, S.M., S.V. Buldyrev, A.L. Goldberger, S. Havlin, R.N. Mantegna, C-K. Peng, M. Simons and H.E. Stanley (1994). Correlation approach to identify coding regions in DNA sequences. *Biophysical Journal* **67**, 64–70.

Peitgen, H.O., H. Jurgens and D. Saupe (1992). *Chaos and Fractals: New Frontiers of Science.* Springer-Verlag, New York.

Peng, C-K., S.V. Buldyrev, A.L. Goldberger, S. Havlin, F. Sciortino, M. Simons and H.E. Stanley (1992). Long-range correlation in nucleotide sequences. *Nature* **356**, 168–170.

Peng, C-K., J.M. Hausdorff, J.E. Mietus, S. Havlin, H.E. Stanley and A.L. Goldberger (1995a). Fractals in physiological control: From heartbeat to gait. In: M.F. Shlesinger, G.M. Zaslavsky and U. Frisch, eds., *Lévy Flights and Related Phenomena in Physics,*

Proceedings of the 1994 International Conference on Lévy Flights. Springer-Verlag, Berlin.

Peng, C-K., S. Havlin, H.E. Stanley and A.L. Goldberger (1995b). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* **5**, 82–87.

Percival, D.P. and P. Guttorp (1994). Long-memory processes, the Allan variance and wavelets. In: E. Foufoula-Georgiou and P. Kumar, eds., *Wavelets in Geophysics.* Academic Press, New York.

Percival, D.P. and A.T. Walden (2000). *Wavelet Methods for Time Series Analysis.* Cambridge University Press, Cambridge.

Robinson, P. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics* **23**, 1630–1661.

Taqqu, M.S., V. Teverovsky and W. Willinger (1995). Estimators for long-range dependence: An empirical study. *Fractals* **3**, 785–798.

Taylor, C.C. and S.J. Taylor (1991). Estimating the dimension of a fractal. *Journal of the Royal Statistical Society B* **53**, 353–364.

Tricot, C. (1995). *Curves and Fractal Dimension.* Springer-Verlag, New York.

Wang, Y. (1997). Fractal function estimation via wavelet shrinkage. *Journal of the Royal Statistical Society B* **59**, 603–613.

Wang, Y., J.E. Cavanaugh and C. Song (1997). Self-similarity index estimation via wavelets for locally self-similar processes. Contributed presentation, International Workshop on Wavelets in Statistics, Duke University, Durham, North Carolina.

Wang, Y., J.E. Cavanaugh and C. Song (2001). Self-similarity index estimation via wavelets for locally self-similar processes. *Journal of Statistical Planning and Inference* **99**, 91–110.

Willinger, W., V. Paxson and M.S. Taqqu (1998). Self-similarity and heavy tails: Structural modeling of network traffic. In: R.J. Adler, R.E. Feldman and M.S. Taqqu, eds., *A Practical Guide to Heavy Tails: Statistical Techniques and Applications.* Birkhäuser, Boston.

Wornell, G.W. and A.V. Oppenheim (1992). Estimation of fractal signals from noisy measurements using wavelets. *IEEE Transactions on Signal Processing* **40**, 611–623.